

Oxidative Damage of Nucleic Acids & its Consequences: Towards an Understanding at an Atomic Level.

A THESIS SUBMITTED FOR THE DEGREE OF DOCTOR OF PHILOSOPHY
IN THE SCHOOL OF CHEMISTRY, FOOD BIOSCIENCES AND PHARMACY.

MEGAN CATHERINE LAMBERT
26831894

March 2023

Declaration

I confirm that this is my own work and the use of all material from other sources has been properly and fully acknowledged.

Abstract

During normal cell metabolism, the regulation of reactive oxygen species (ROS) is important to keep levels appropriate to maintain cell function. However, during oxidative stress, when ROS production increases and the cell's ability to neutralise ROS is diminished, many cellular components can be damaged. Oxidative DNA damage induced by ROS, the formation of which can be catalysed by transition metals, has been reported as frequent in an oxidative stress environment.

Transition metals bind to DNA via the phosphate backbone and the endocyclic atoms of nucleic acids e.g. at the N7 atoms of purine bases. They have also been shown to catalyse the formation of reactive oxygen species, including hydroxyl radicals HO•, and may therefore act as hotspots for oxidative damage with nearby bases. It is unknown if catalysis occurring at these spots would damage the closest base or if the bound transition metal would protect the base it is bound to from damage.

Using X-ray crystallography, a common technique used to study nucleic-acid structures at an atomic scale, the work presented in this thesis describes the development of a new method to identify oxidative damage hotspots in nucleic-acid structures. The objective is to treat nucleic acid crystals, containing ordered bound metal sites, with hydrogen peroxide to induce oxidative damage. The effects of oxidative damage on nucleic acid structures were also investigated in solution using circular dichroism (CD) spectroscopy.

Two approaches are used to add transition metals to crystal systems containing nucleic acids. These are cocrystallisation and soaking - the advantages and disadvantages of these two methods in the context of this work were carefully examined. Soaking was considered as a more suitable approach, as cocrystallisation was found to change the structure and cause crosslinking which reduced transition-metal availability to hydrogen peroxide. Applying this method, the binding preferences of transition metals such as copper and iron were investigated. Two transition metals, copper and iron, were found to bind differently to the same DNA structure, a result that could explain why oxidative damage differs between systems containing copper or iron. Furthermore, the same system was used to soak DNA crystals with hydrogen peroxide for the first time. No damage was identified but ordered hydrogen peroxide molecules bound to

copper (II) ions were described, giving insight into how the Fenton reaction might progress when copper (II) is bound to the DNA. Finally, using CD, a G-quadruplex forming sequence present in the promoter region of PSEN2, a gene implicated in Alzheimer's disease, was found to undergo conformational changes in an oxidative stress environment. The structural effects of the presence of more than four tracts of guanines in the sequence were investigated, using base substitutions to suppress quadruplex formation and mimic oxidative damage. This gave insight into the potential equilibrium between all the potential G-quadruplex topologies in this system and how oxidative damage can influence topological changes.

This method is a starting point to examine more sophisticated nucleic-acid systems such as G-quadruplexes. Whilst oxidative damage could not be directly visualised in the crystal structures, the use of microcrystals combined with micro-focused X-ray beams at synchrotrons or XFEL sources may lead to visualisation in the future.

Acknowledgements

When I started on this project four years ago, nothing could have prepared me for a worldwide pandemic and the disruption it caused. These last three years have been hard for everyone, and I am no exception. This work could not have been made possible without a lot of guidance and support from many people.

First and foremost, I would like to express my gratitude to my supervisors, Dr James Hall, Dr John Brazier, and Dr Juan Sanchez-Weatherby, for their invaluable patience and encouragement throughout my doctoral studies. Without them, I would not have made it this far, and for this, I am sincerely grateful. Their genuine enthusiasm for science has made working on this project an immensely enjoyable experience.

I would also like to thank Professor Christine Cardin, Professor David Cardin, Dr Kane McQuaid, Dr Ben Pages and Dr Stuart Malthus from the University of Reading, for their guidance at the start of my PhD years. I would also like to thank Nick Spencer for his help particularly with the in-house diffractometer and Ahmad Abdullrahman for his cheerfulness which has never failed to put me in a good mood.

Also, I would like to thank everyone at Diamond Light Source, and especially the teams on the VMXI, I03, I04, and B23 beamlines, as well as the team at the Research Complex at Harwell, for all the opportunities I've had during user visits. I truly enjoyed working with all these passionate and wonderful people.

I would also like to thank Professor Tom Brown, Professor James Tucker and Dr Sarah Allman, who welcomed me into their teams and opened the doors to their labs when I needed it most.

On another note, I would like to thank my psychiatrist Dr De Waal for recognising my battles, the University of Reading for their incredible support, and especially Dr James Hall who spared no effort to support me after my diagnosis. This means the world to me.

Finally, I would like to thank my partner, Matthew Greenleaf, my friend, Amélie Boullais, and my gecko, Coconut, who have been my rock during write up, as well as my family and friends, and everyone else who has supported and encouraged me during this chapter of my life.

Table of Contents

ABSTRACT	II
ACKNOWLEDGEMENTS	IV
ABBREVIATIONS USED	VIII
CHAPTER 1 - AN INTRODUCTION TO DNA MODIFICATIONS AND DNA DAMAGE: THE SUBSEQUENT EFFECTS ON DNA STRUCTURE	1
1.1. INTRODUCTION	1
1.2. NATIVE DNA STRUCTURES	2
1.2.1. <i>Base-pairing</i>	4
1.2.2. <i>Secondary structure of nucleic acids</i>	7
1.2.3. <i>Higher-order structure of nucleic acids</i>	9
1.3. OVERVIEW OF NUCLEIC ACID BASE MODIFICATIONS AND THEIR STRUCTURAL CONSEQUENCES	11
1.3.1. <i>Epigenetic modifications</i>	11
1.3.2. <i>Damage of nucleic acid bases</i>	14
1.4. STRUCTURAL CONSEQUENCES OF BASE MODIFICATIONS	18
1.4.1. <i>Effects of epigenetic modifications on the overall structure</i>	18
1.4.2. <i>Effects of damaged bases on the overall structure of nucleic acids</i>	24
1.5. INCORPORATION OF TRANSITION METALS INTO NUCLEIC ACID STRUCTURES: HOW THE IDENTIFICATION OF TRANSITION METAL CENTRES CAN HELP IDENTIFY POTENTIAL DAMAGE CENTRES	28
1.6. AIMS	30
1.7. REFERENCES	33
CHAPTER 2 - CO-CRYSTALLIZATION VS SOAKING - A COMPARISON OF METHODOLOGIES FOR OBTAINING ORDERED METAL SITES IN CRYSTALS WITH MINIMUM STRUCTURAL PERTURBATION	50
2.1. INTRODUCTION	50
2.2. MATERIAL AND METHODS	55
2.2.1. <i>DNA preparation</i>	55
2.2.2. <i>Crystallisation and metal-soaking</i>	55
2.2.3. <i>Data collection and processing</i>	56
2.2.4. <i>Structural and packing analysis</i>	59
2.3. RESULTS	59
2.3.1. <i>Overall structure analysis - soaking experiment</i>	59
2.3.2. <i>Comparison to results from cocrystallisation</i>	62
2.3.3. <i>Transition metal coordination analysis and solvent channel availability</i>	64
2.4. DISCUSSION	67
2.5. REFERENCES	72

CHAPTER 3 - DAMAGE FOLLOWING FENTON-LIKE COPPER REACTION INFLUENCES	
COPPER (II) IONS BINDING PREFERENCES TO DNA STRUCTURE	77
3.1. INTRODUCTION	77
3.2. MATERIAL AND METHODS	79
3.2.1. DNA preparation.....	79
3.2.2. Crystallisation of <i>d(GCATGCT)</i> and <i>d((oxoG)CATGCT)</i> and soaking with copper (II) chloride	79
3.2.3. Data collection and data analysis for a crystal of <i>d(GCATGCT)</i> with copper (II) chloride and a crystal of native <i>d((oxoG)CATGCT)</i>	79
3.2.4. Structural analysis.....	82
3.3. RESULTS	82
3.3.1. <i>oxoG</i> substitution in a non-classical quadruplex structure	82
3.3.2. Copper (II) ions coordination analysis after soaking of native crystals.....	86
3.4. DISCUSSION	91
3.4.1. Effects of oxidative damage on metal ions coordination	91
3.5. REFERENCES	96
CHAPTER 4 - IRON (II) BINDING PREFERENCES TO DNA STRUCTURE – OBSERVING METAL	
COORDINATION BEFORE OXIDATIVE DAMAGE.....	100
4.1. INTRODUCTION	100
4.2. MATERIAL AND METHODS	102
4.2.1. DNA preparation.....	102
4.2.2. Crystallisation and soaking experiment	103
4.2.3. Data collection and processing	104
4.2.4. Structural and packing analysis.	106
4.3. RESULTS	107
4.3.1. DNA hexamer: overall structure and helical parameters analysis	107
4.3.2. Cu^{2+} and Fe^{2+} ions coordination analysis.....	111
4.4. DISCUSSION	116
4.5. REFERENCES	120
CHAPTER 5 - STRUCTURAL ANALYSIS OF PEROXIDE-SOAKED DNA CRYSTALS CONTAINING	
ORDERED COPPER BINDING SITES: TOWARDS UNDERSTANDING OXIDATIVE DAMAGE AT	
THE ATOMIC SCALE.....	125
5.1. INTRODUCTION	125
5.2. MATERIAL AND METHODS	127
5.2.1. DNA preparation.....	127
5.2.2. Crystallisation of <i>d(CGCGCG)</i>	127
5.2.3. Time-dependent in-crystal Fenton reaction and flash-freezing.....	127
5.2.4. Synchrotron data collection and processing	128

5.2.5. <i>In-house multi dataset collection at room-temperature from a peroxide-soaked crystal as a function of time</i>	130
5.2.6. <i>Structural and packing analysis</i>	133
5.3. RESULTS	134
5.3.1. <i>Peroxide-soaked d(CGCGCG) crystals: 100 K collection</i>	134
5.3.2. <i>Peroxide-soaked d(CGCGCG) crystals: room temperature collection</i>	140
5.4. DISCUSSION	146
5.5. REFERENCES	149
CHAPTER 6 - CONFORMATIONAL CHANGES MEDIATED BY A FENTON-LIKE COPPER REACTION IN A G-QUADRUPLEX FORMING SEQUENCE	152
6.1. INTRODUCTION	152
6.2. MATERIAL AND METHODS.....	158
6.2.1. <i>DNA and samples preparation</i>	158
6.2.2. <i>Quadruplexes formation and melting experiments</i>	158
6.2.3. <i>Fenton reaction monitoring using Circular Dichroism</i>	159
6.2.4. <i>Sample preparation for analysis</i>	159
6.2.5. <i>Preparation of denaturing polyacrylamide gels</i>	159
6.3. RESULTS	160
6.3.1. <i>Confirmation of G-quadruplex formation by circular dichroism</i>	160
6.3.2. <i>Effects of oxidative damage on PSEN2 topology after thermal denaturation</i>	161
6.3.3. <i>Effects of G-tract substitution on PSEN2 topology</i>	162
6.3.4. <i>Identifying strand breaks after the Fenton reaction</i>	165
6.4. DISCUSSION	166
6.5. REFERENCES	170
SUMMARY AND FUTURE WORK	175
REFERENCES.....	179
SUPPLEMENTARY INFORMATION	180

Abbreviations Used

A	Adenine
CD	Circular dichroism
CCP4	Collaborative computational project number 4
C	Cytosine
DNA	Deoxyribonucleic acid
T _m	DNA melting temperature
Ds	double-stranded DNA
G	Guanine
N	Nitrogen
oxoG	8-oxoguanine
O	Oxygen
RNA	Ribonucleic acid
Pu	Purine
Ss	single-stranded DNA
T	Thymine
Py	Pyrimidine

Chapter 1 - An Introduction to DNA Modifications and DNA Damage: The Subsequent Effects on DNA Structure

1.1. Introduction

Nucleic acids, deoxyribonucleic acid (DNA) and ribonucleic acid (RNA), are some of the most important biomacromolecules in living organisms and play a fundamental role in storing, transferring, and encoding genetic information. Nucleic acids were first discovered in 1871 by Friedrich Miescher after he made the DNA from leucocytes (white blood cells) precipitate and were generally accepted as genetic material in the late 1800's¹. Yet, the popularity of this hypothesis declined in the 1920's when nucleic acid structure started being better understood, and proteins were seen as more diverse in type and complexity². Nevertheless, it was not until Griffith discovered genetic "transformation" between different strains of the bacterium pneumococcus in 1928 and the identification of DNA as the substance responsible for this phenomenon by Avery, MacLeod, and McCarty in 1944, that the idea nucleic acids are genetic material was made widely popular again².

Nowadays, the correlation between the structure of nucleic acids and their functions is well established³⁻⁵. Nucleic acids possess both stability and flexibility, two characteristics essential for storing genetic information while still enabling evolution^{6,7}. However, many mechanisms surrounding nucleic acids and their roles are yet to be fully understood and particularly how nucleic acid damage and modifications takes part in both normal functioning and breakdown⁸.

1.2. Native DNA structures

Nucleic acids are biopolymers made of monomeric units called nucleotides, linearly and covalently bonded to form a larger structure. All units are constructed from three components: a heterocyclic base containing a nitrogen, a pentose sugar and a phosphate group^{3,9}. Bases fall into two categories: purines and pyrimidines. Purines (R) are bicyclic and consist of two fused five and six-membered rings. The two standard forms found in both DNA and RNA are adenine (A) and guanine (G). Pyrimidines (Y) are monocyclic and contains six atoms in the ring. The two standard forms found in DNA are cytosine (C) and thymine (T). However, thymine is not present in RNA and is replaced by uracil (U). Uracil differs from thymine by the absence of a methyl group at the fifth position of its ring (Figure 1.1).

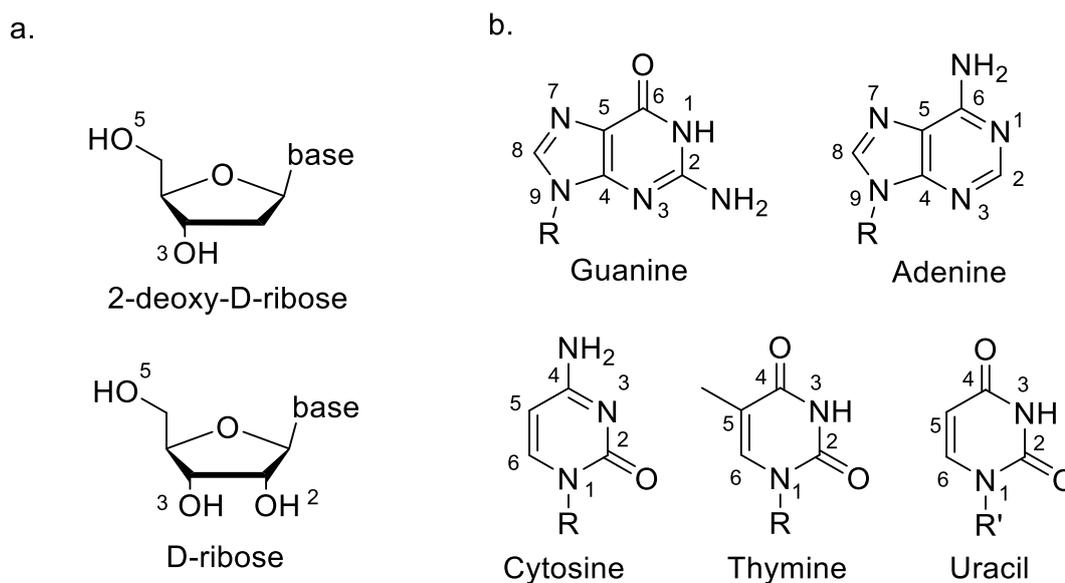


Figure 1.1: a. The two pentose sugars of nucleic acids. b. The five canonical bases of nucleic acids. The R group on the bases represent where the base joins to the sugar ring. The nucleotides are shown with standard numbering convention. Groups or atoms bonded to a ring atom have the same numbers.

In RNA, the pentose sugar is D-ribose forming a five-membered furanose ring while in DNA, the pentose is 2-deoxy-D-ribose. The base is connected to the sugar via a β -glycosidic bond between the C1 position of the ring and either the N1 or N9 of the base¹⁰. The compact nature of the ring and the interactions between its non-bonded atoms make it pucker out of plane and

can be described by the displacement of the carbon out of plane¹¹. Pucker assignment is based on the pseudo-rotation phase angle, P, defined as:

$$\tan P = \frac{(\theta_2 + \theta_4) - (\theta_1 + \theta_3)}{2\theta_0(\sin 36 + \sin 72)}$$

All possible sugar pucker conformations along with their P values are depicted in Figure 1.2. The main two conformations are *C2'-endo* and *C3'-endo*¹² as illustrated in Figure 1.2. Finally, each nucleotide is joined to the next at the 3'-hydroxyl group, which forms a bond with the 5'-hydroxyl group of an adjacent nucleotide to form a phosphate diester, forming the phosphate backbone.

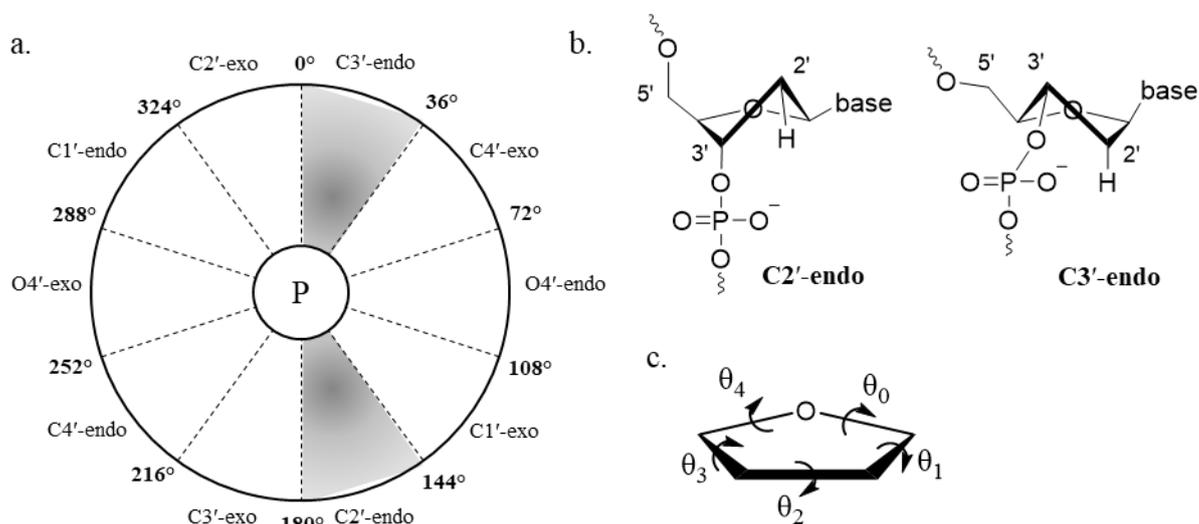


Figure 1.2: a. A diagram of the phase angle of pseudo-rotation P. b. The two main sugar conformations illustrated. c. Designation of the different sugar pucker torsion angles. The figures were made using Powerpoint and Chemdraw.

1.2.1. Base-pairing

Supported by Chargaff's work from 1950¹³ which showed 1:1 ratios of adenine:thymine and guanine:cytosine in DNA, and thanks to Franklin's diffraction pattern, Watson and Crick proposed in 1953 a double-helical structure with complementary hydrogen-bonded base-pairs¹⁴. These pairings became known as the Watson-Crick pairing. Two hydrogen bonds between the adenosine and the thymine are present in a Watson-Crick A:T pair while three hydrogens bonds are present between the cytosine and the guanine in a Watson-Crick G:C pair (Figure 1.3).

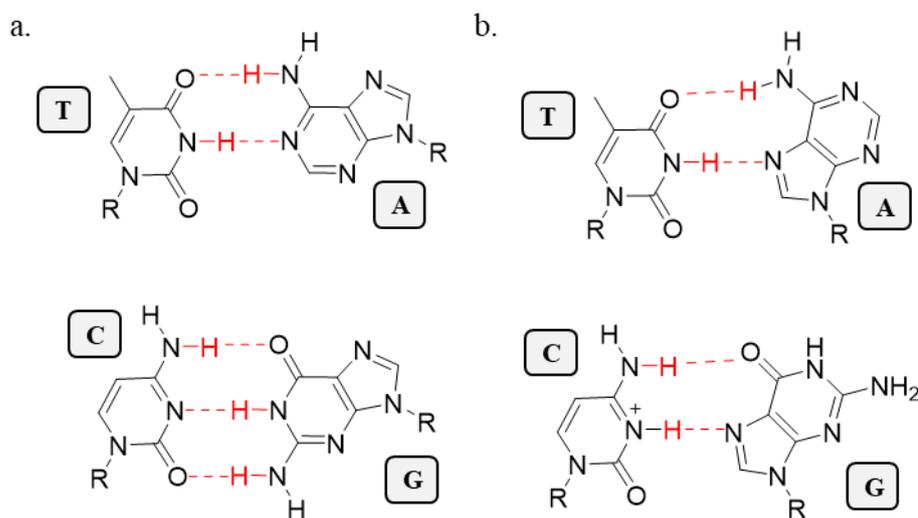


Figure 1.3: a. Watson-Crick pairing between an adenine and a thymine (top); and a guanine and a cytosine (bottom). b. Hoogsteen base pairing between an adenine and a thymine (top); and a guanine and a cytosine (bottom). Hydrogen bonds are shown as red dashed lines.

While these pairings are dominant, non-canonical base-pairings can also occur. In 1963, Karl Hoogsteen discovered A:T pairs with a different geometry to Watson-Crick pairs¹⁵. Later in 1966, Crick published the “Wobble Hypothesis” to explain possible non-canonical base pairings between tRNA and mRNA¹⁶. Later, G:C pairs were also discovered to have an alternative geometry to Watson-Crick pairs. Hydrogen donors and acceptors are available on other edges of the different bases and can allow to form either polar hydrogen bonds (N-H***O/N or O-H***O/N) or non-polar hydrogen bonds (C-H***O/N). All the three different edges where hydrogen sites are present in both pyrimidines and purines were named: the Watson-Crick edge, the Hoogsteen edge and the sugar edge (Figure 1.4). Non-canonical base-

pairings like Hoogsteen base pairs are found in both DNA, as for example in triplexes¹⁷ and quadruplexes¹⁸, and RNA, as in the structure of the tRNA¹⁹.

Pairings can also be described as *cis* or *trans* depending on the sugar's orientation²⁰ (Figure 1.4). By combining cis/trans and different edges, a dozen pairing combinations between bases can be obtained. Base-pairing also follows rigid-body base-pair parameters that are linked to the direction the hydrogen bonds and are summarised in Figure 1.5.

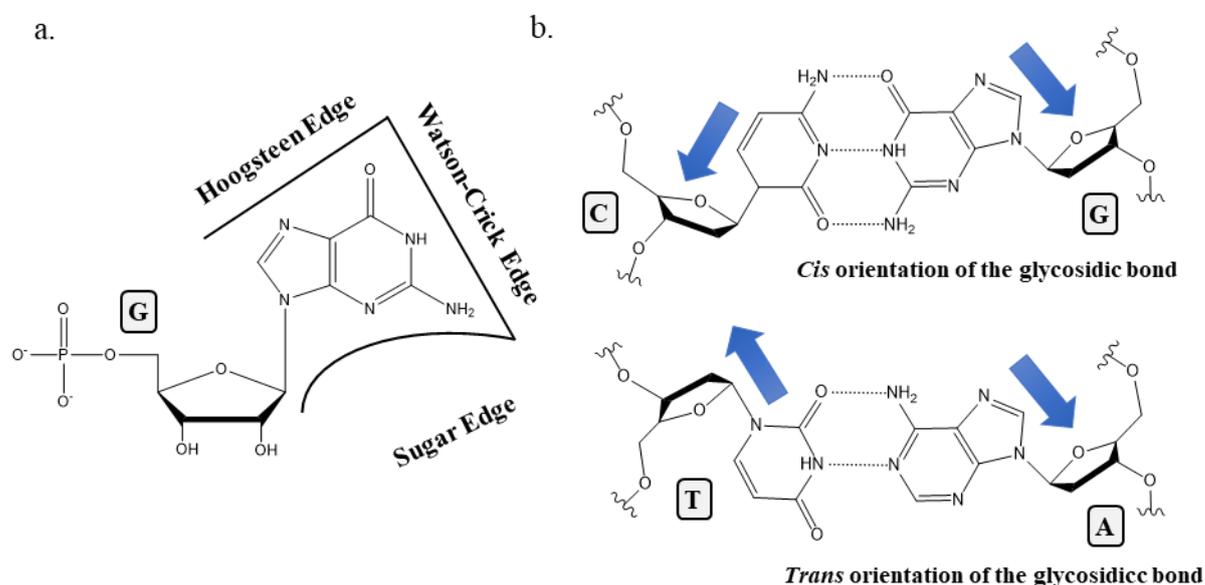


Figure 1.4: a. The Watson–Crick, Hoogsteen and Sugar edges on a guanosine nucleotide. b. *Cis* vs *Trans* orientation of the glycosidic bond, as illustrated with a C:G base pair (top) and T:A base pair (bottom).

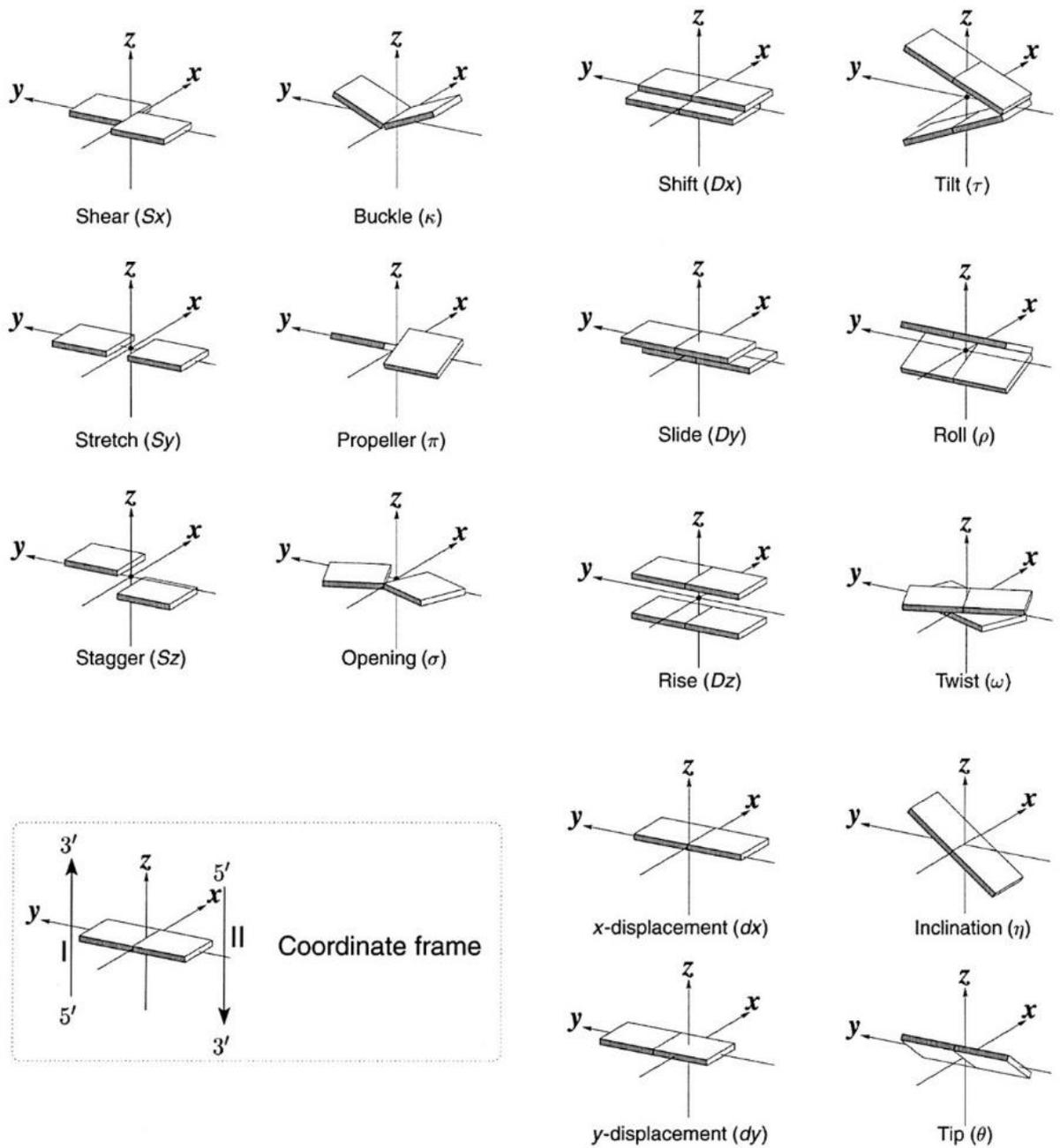


Figure 1.5: The individual base pair and base pair step parameters. Each base is represented by a rectangular block. Copyright 2003 Oxford University Press²¹.

1.2.2. Secondary structure of nucleic acids

From fibre diffraction studies, two of the three most common duplex structures of DNA, B-DNA and A-DNA, were characterised thanks to the work of Watson, Crick, Franklin, and Gosling in the early 50's^{14,22,23}. Z-DNA, the third most common duplex structure of DNA, was characterised a couple decades later by Wang *et al* (Figure 1.6)²⁴. The B-DNA form is observed at high humidity and commonly corresponds to the form found in physiological conditions, though it will convert to the A-DNA form once the humidity is lowered²⁵ or at high salt concentrations. The helix is right-handed and makes a complete turn every ten bases. Its phosphate backbone forms two grooves, a minor groove with a width of 5.7 Å and depth of 7.5 Å and a major groove 11.7 Å wide and 8.8 Å deep²⁶. The A-DNA form is also right-handed, makes a complete turn every eleven bases, with a minor groove 11.1 Å wide and 2.6 Å deep and a major groove 13.0 Å wide and 2.2 Å deep. RNA also adopts the A form when it is in a double-helical structure. Finally, unlike its A-DNA form and B-DNA counterpart, the Z-DNA helix is left-handed. This contains a high number of CG repeats and is very different to the other two forms. All three duplexes are represented in Figure 1.6. All parameters are summarised in Table 1.1^{3,9,26}.

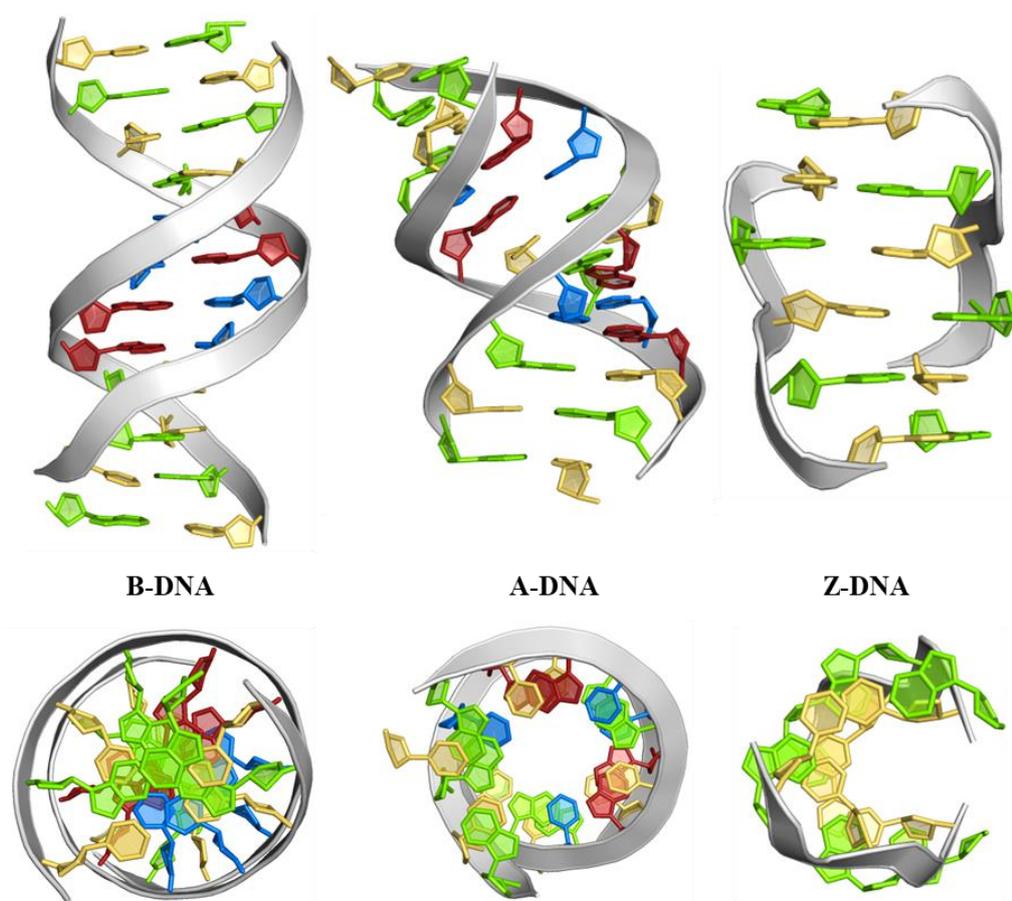


Figure 1.6: Views of the three conformations of double helices DNA (top) along with the view from the top (bottom). Backbone chain is in light grey. Nucleotides are coloured following convention, guanine bases in green, adenine bases in red and thymine bases in blue. The structures shown are taken from entries 117D²⁷ (A-DNA), 1BNA²⁸ (B-DNA) and 3P4J²⁹ (Z-DNA) in the Protein Data Bank³⁰ (PDB).

Table 1.1: Average parameters of all three types of duplex structures from X-ray analysis^{3,9,26}.

	A-DNA	B-DNA	Z-DNA
Helix sense	Right-handed	Right-handed	Left-handed
Residues per turn	11	10	12
Twist per bp	32.7	36	-9, -51
Displacement bp (Å/Å)	4.5	-0.2 to -1.8	-2 to -3
Rise per bp/Å	2.56	3.3-3.4	3.7
Base tilt (°)	20	-6	-7
Sugar pucker	C3'-endo	C2'-endo	C3'-endo (<i>syn</i>)
Groove width (minor) (Å)	11.1	5.7	2.0
Groove depth (minor) (Å)	2.6	7.5	8.8
Groove width (major) (Å)	13.0	11.7	3.7
Groove depth (major) (Å)	2.2	8.8	13.8

1.2.3. Higher-order structure of nucleic acids

Even though double-stranded structures are considered to be the most common structural form of DNA, it has long been recognised that DNA and RNA can also form other structures including triple-stranded helices^{17,31}, hairpin loops^{32,33}, junctions^{34,35} and four-stranded quadruplexes^{18,36}. Some of these structures have been widely studied for their effect on genetic instability^{37,38} and link to biological processes such as DNA repair and some neurodegenerative diseases such as Alzheimer's disease or Parkinson's disease³⁹. G-rich sequences can form four-stranded quadruplexes known as G-quadruplexes³⁶ and their counterpart, C-rich sequences, can also form four-stranded structures known as i-motifs⁴⁰.

Found in the telomeric and gene promoter regions of the genome, G-quadruplexes form from sequences containing four tracts of three to four guanines⁴¹. The cyclic arrangements between the four guanines are stabilised by Hoogsteen interactions, which are themselves stacked on top of each other to form the four-stranded quadruplex. This formation is stabilised by the presence of sodium and/or potassium cations in the channel, as can be seen in Figure 1.7. G-quadruplexes can be either unimolecular, bimolecular or tetramolecular, depending on the number of strands involved in their formation. Quadruplex topologies are classified in three different sub-groups: parallel, antiparallel and hybrid. Bases not directly involved in the base G-tetrads are part of the 'loop' region, linking consecutive G-tracts, and are categorised in three types: propeller, lateral, and diagonal⁴².

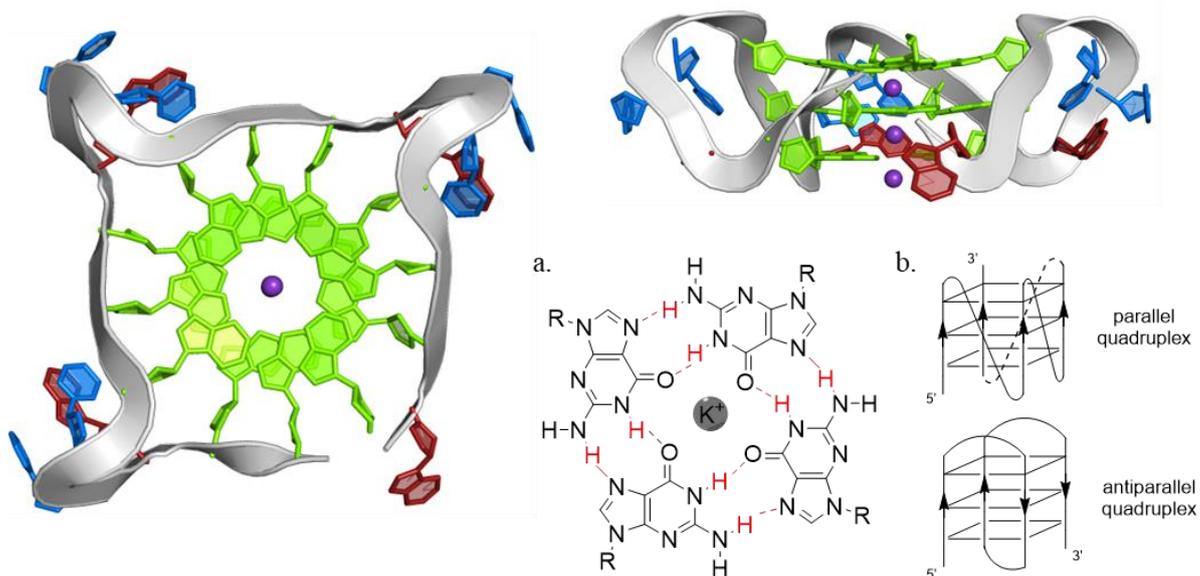


Figure 1.7: Crystal structure of the human telomeric G-quadruplex (PDB ID: 1KF1⁴³). The backbone chain is in light grey. Nucleotides are coloured following convention, guanine bases in green, adenine bases in red and thymine bases in blue. The purple spheres in the channel are potassium cations which stabilise the structure. a. Schematic showing how a G-quadruplex is stabilized by Hoogsteen hydrogen bonding with a potassium cation in the channel. b. Some topologies of intramolecular quadruplexes that have been reported.

Characterised *in vitro* in 1993 with the DNA sequence d(TCCCCC), i-motifs have been identified to form four-stranded quadruplexes under acidic conditions⁴⁰. For quite some time, this led to a discourse between researchers regarding its biological relevance *in vivo*. Since then, studies have shown that i-motifs can form at physiological pH, under the condition of molecular crowding. More recently, a study published in 2018 has demonstrated that i-motif structures can form within the nuclei of human cells, finally showing that this structure can exist in a biological environment⁴⁴.

Unlike G-quadruplexes, i-motifs are parallel-stranded duplexes held together through intercalated base pairs⁴⁵. This leads to a structure containing two major grooves and two minor grooves and due to its spatial arrangement, two types of topologies have been identified, 3'E and 5'E (Figure 1.8)⁴⁶. Intra- and intermolecular i-motifs have also been reported, and similarly to G-quadruplexes, tetramolecular and bimolecular i-motifs have also been identified.

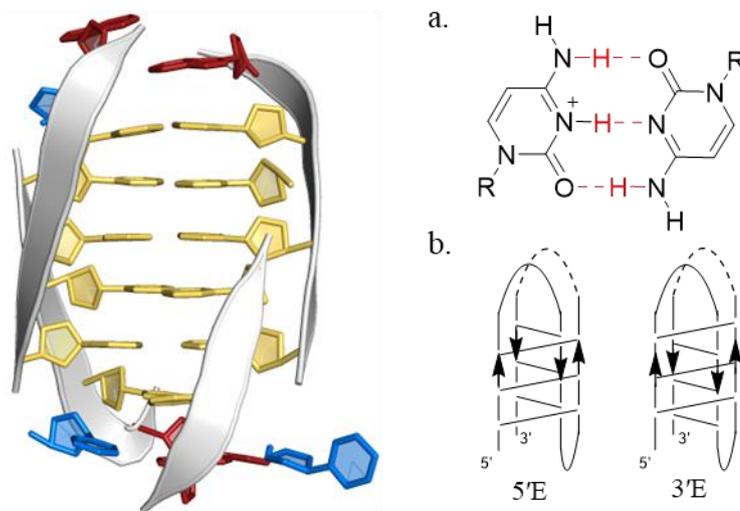


Figure 1.8: Crystal structure of a four-stranded i-motif (PDB ID: 1CN0⁴⁵). The backbone chain is in white. Nucleotides are coloured following convention, cytosine bases in yellow, adenine bases in red and thymine bases in blue. a. Cytosine-cytosine pair stabilized by Watson-Crick hydrogen bonding, here represented by red dashed lines. b. Two topologies of intramolecular i-motifs that have been reported.

1.3. Overview of nucleic acid base modifications and their structural consequences

In the normal cellular environment, nucleotides can be chemically modified, which will cause either epigenetic modification⁴⁷, affecting gene expression without altering the sequence, or DNA lesions, by damaging DNA which will result in changes in the coding properties if not repaired^{48,49}. Either way, DNA is subject to modifications on a molecular level which will impact its functionality. In the last decades, biologists have been studying biological processes and chemical reactions that have been leading to nucleotides modifications under physiological conditions, towards a comprehensive understanding of their effects in health and disease⁵⁰.

1.3.1. Epigenetic modifications

The term “epigenetic”, meaning “on top of changes in genetic sequence”, is a term first introduced by Conrad Waddington in the early 40’s^{51,52}. And although epigenetic modifications can be categorised into three groups: DNA methylation, histone modifications and nucleosome positioning⁵⁰, in this work, only the first group will be presented. Methylation of cytosine at the

fifth carbon by enzymes has been identified as the most common modification in DNA linked to epigenetics, so much that 5-methylcytosine (5mC) has been dubbed the “fifth base” of DNA⁵³. As much as 5 % of cytosines has been detected to be methylated in the human brain⁵⁴, and as high as 80 % of CpG dinucleotides and CpG islands present in the whole genome have been shown to be methylated^{55–57}. Other DNA modifications such as 5-hydroxymethylcytosine (5hmC), 5-formylcytosine (5fC), 5-carboxycytosine (5caC), which were discovered in the human genome late 2000s, early 2010s^{58,59}, or 6-methyladenine (6mA), 6-hydroxymethyladenine (6hmA) or 6-formyladenine (6fA) have also been shown to be present in DNA in smaller quantity^{47,60,61}. Regulation of genes via methylation is a dynamic process which involves active and passive modifications; however, all mechanisms are not well known.

Methylation of DNA in mammals is catalysed by a family of enzymes called DNA methyltransferases (DNMT). They have been shown to catalyse three different chemical reactions: methylation of the sixth position of adenines, methylation of the fourth position of cytosines and methylation of the fifth position in cytosines^{62–64}, thus forming three different groups of DNMT. The demethylation pathway that has been the most studied involves 5-methylcytosine and its oxidation products. Demethylation of DNA is actively promoted by ten-eleven translocation (TET) enzymes. The TET enzymes catalyse hydroxylation of 5mC to 5hmC, base which can then be further oxidised to 5caC and 5fC⁵⁹. The thymine-DNA glycosylase (TDG) was reported to be able to remove 5fC and 5caC efficiently, excising them and initiating the base excision repair (BER) pathway, which replaces the abasic site with an unmodified cytosine⁶⁵ (Figure 1.9). During the demethylation process, TET enzymes have also been shown recently to be able to be recruited by other enzymes and one of these enzymes, the oxoguanine glycosylase (OGG1), can initiate the demethylation pathway in the presence of 5mCp-8-OHdG dinucleotide⁶⁶. The presence of 6mA in eukaryotic cells however has been controversial as it was believed it only had an important role in bacteria, up until recently when it was discovered in small eukaryotic organisms^{67,68}. The methylation/demethylation pathway of 6mA is believed to be similar to 5mC, with the demethylation initiated by enzymes from the AlkB protein family⁶⁹ (Figure 1.10).

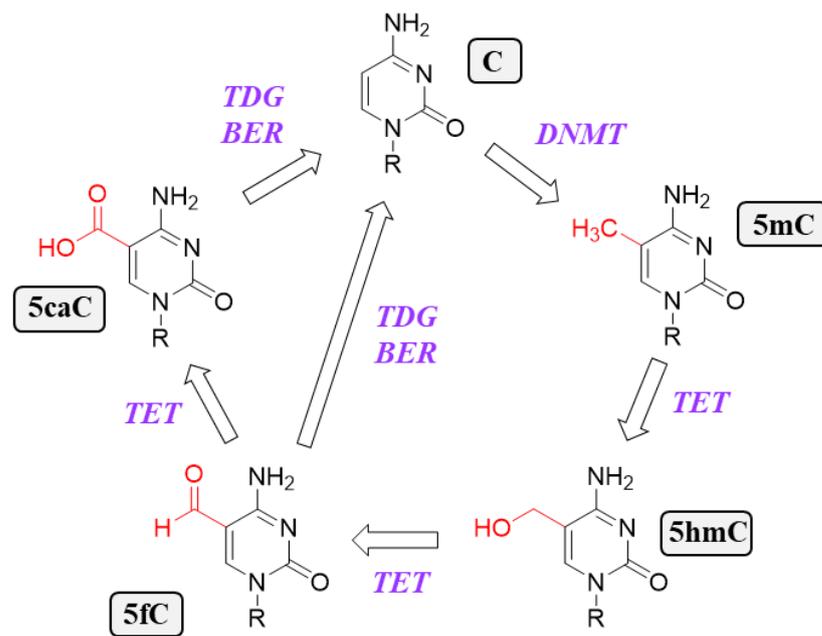


Figure 1.9: The cytosine methylation/demethylation pathway. Cytosine (C) modifications to 5-methylcytosine (5mC) and 5-hydroxymethylcytosine (5hmC) are mediated by the DNA methyltransferase (DNMT) and ten-eleven translocation (TET) enzymes (purple). The oxidation products 5-formylcytosine (5fC) and 5-carboxycytosine (5caC) are then recognised by the thymine-DNA glycosylase (TDG) and excised, to be replaced by unmodified cytosines through the base excision repair (BER) pathway. Modifications are highlighted in red.

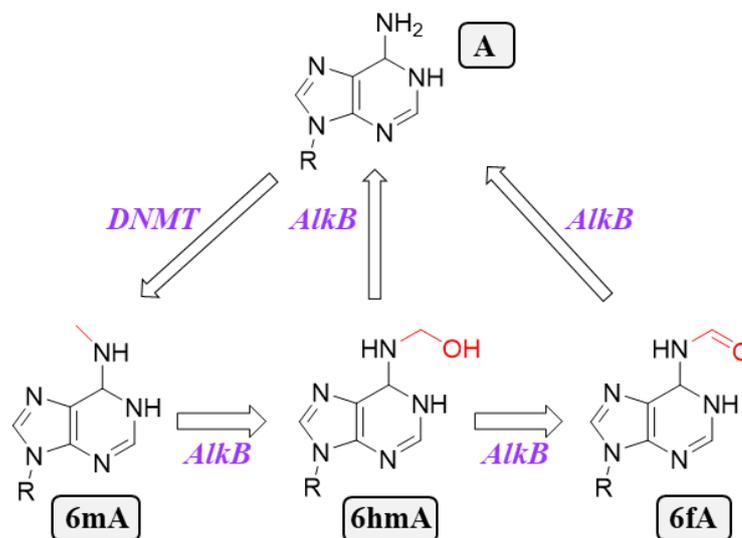


Figure 1.10: The adenine methylation/demethylation pathway in prokaryotes. Modifications are mediated by the DNA methyltransferase (DNMT) and AlkB enzymes (purple). Demethylation to adenine is initiated by AlkB after recognition of 6-hydroxymethyladenine and 6-formyladenine. Modifications are highlighted in red.

1.3.2. Damage of nucleic acid bases

Alkylated DNA lesions can originate from exogenous sources (chemical agents like nitrogen mustards⁷⁰ and chloroethylating agents as anticancer drugs⁷¹, or other sources like tobacco smoke⁷² and fuel combustion⁷³) or from endogenous metabolites (from gut bacteria activity⁷⁴ or from lipid peroxidation⁷⁵). Nitrogen and oxygen atoms are mostly targeted by alkylating agents in nucleic acids because of their nucleophilic nature and high negative electrostatic potential⁷⁶, although differences in reactivity depends on the position and whether the nucleic acid is paired or not. However, N-methyl adducts still comprise the majority of alkylation products. Some of the most studied methylating agents like N-methyl-N-nitrosourea (MNU) and N-methyl-N'-nitro-N-nitrosoguanidine (MNNG) react with a SN1 type mechanism, while methyl methane sulfonate (MMS) and dimethyl sulfate (DMS) react with a SN2 type mechanism.

Among all the lesions that have been identified caused by common methylating agents, the most abundant alkylation product in double stranded DNA was found to be N7-methylguanine (7mG), with for example 67 % and 82 % of the total lesions caused by MNNG and MMS respectively⁷⁷. N3-methyladenine (3mA) has also been determined to be the second most abundant lesions, accounting for around 11-12 % of total lesions. MNNG and MMS also produce 7 % and 0.3 % O6-methylguanine (6mG) respectively. Other products include N1-methyladenine (1mA), N7-methyladenine (7mA), N1-methylguanine (1mG), N3-guanine (3mG), N3-cytosine (3mC), O2-cytosine (2mO), N3-thymine (3mT), O2-thymine (2mT), and O4-thymine (4mT) and accounted for less than 5 %⁷⁷. The possible sites of alkylation in DNA bases have been illustrated in Figure 1.11.

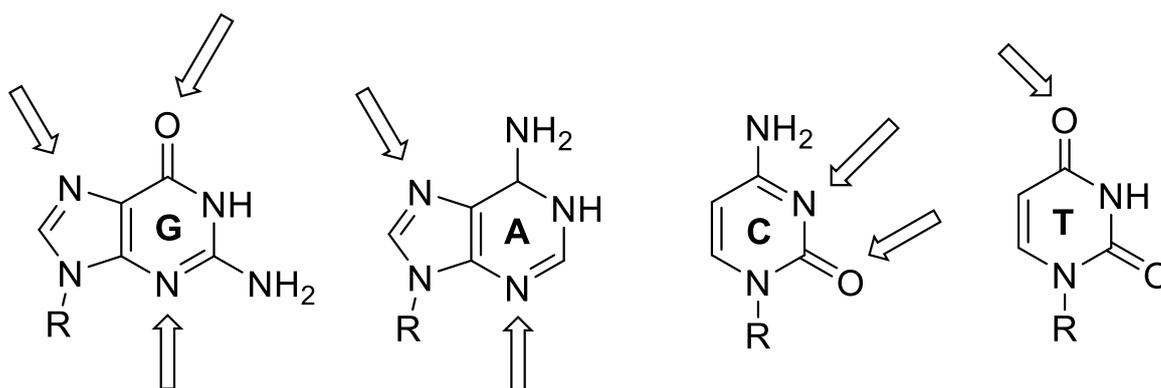


Figure 1.11: Possible sites of alkylation of DNA bases.

Oxidative stress, which is the imbalance between reactive oxygen species (ROS) and antioxidant agents⁷⁸, often leads to oxidation of DNA. It has been implicated in many diseases (neurodegenerative conditions⁷⁹, cardiovascular disorders⁸⁰, cancers⁸⁰, lung diseases⁸¹) and has been widely investigated. Reactive oxygen species can originate from exogenous (cigarette smoke⁸², pollutant⁸³, UV light⁸⁴, ionizing agents⁸⁵) or endogenous (side products of mitochondrial function⁸⁶ or during enzyme reactions) species. Different chemical processes are involved with both sources to form diverse free radicals. For instance, hydroxyl radicals (HO•), which are highly reactive and have been shown to react with organic molecules, can be produced through a variety of different paths. For example, the Fenton reaction is a key pathway which generates hydroxyl radicals by reaction between transition metals and hydrogen peroxide^{87,88}. This pathway is of great importance in Alzheimer's disease as failed regulation of iron leads to abnormal levels of the metal in patients⁸⁹. Formation of hydroxyl radicals can also be generated via the Haber-Weiss reaction between hydrogen peroxide and superoxide⁹⁰ (O₂-•). Additionally, peroxynitrite decomposition⁹¹ and X-ray radiolysis of water⁹² are other ways of producing hydroxyl radicals. Other radicals such as alkylperoxyl radicals from lipids or amino acids⁹³ and superoxide radicals generated by oxidases⁹⁴ were also reported to damage the DNA.

DNA oxidation can occur in different positions on the DNA strands and result in altered nucleosides. Resulting modifications of nucleosides have been described in the literature, especially those which occur to guanine due to its low oxidation potential⁹⁵ making it the most susceptible to modification. Three major positions have been shown to be targeted and reported

to be on the C4, C5 and C8 positions of the purine. One of the major products of guanine oxidation in the C8 position is 8-oxo-7,8-dihydroguanine, also known as oxoguanine (oxoG), and this has been used as a biomarker of oxidative stress *in vivo*⁹⁶ (Figure 1.12). This product has been reported to lead to the hydantoin lesions spiroiminodihydantoin (dSp) and guanidinohydantoin (dGh) by reaction on the C5 position. Recently 5-carboxamido-5-formamido-2-imino-hydantoin (d2Ih), a C5 position product, has been reported in yield similar to oxoG after DNA oxidation via Fenton-like reaction. C5 position products such as imidazole and oxazolone were also described. Additionally, chemical modifications of adenine were reported to those of guanine though less frequent due to a higher redox potential⁹⁷ and include 8-oxoadenine (8-oxoA) and 2-oxoadenine (2-oxoA). Modifications to cytosine and thymine have also been reported. For example, the oxidation products have also been shown to be a product of Fenton chemistry⁹⁸. Addition on the 5,6 double bond of cytosine by ROS can occur, though with a preference for C5, leading to the 5-hydroxycytosine (5hC), 5-hydroxyuracil (5hU), 5,6-dihydroxy-5,6-dihydrouracil (Ura-Gly), 5-hydroxyhydantoin (Hyd-Ura) and 1-carbamoyl-4,5-dihydroxy-2-oxoimidazolidine (Imid-Cyt) products⁹⁹⁻¹⁰¹. Relative yields of each product were described as different while exposed to ionizing radiation and Fenton-like reagents although both methods produce hydroxyl radicals. It was also reported that thymine oxidation could occur and that this was less prevalent than cytosine oxidation. 5,6-dihydroxy-5,6-dihydrothymine (Thy-Gly), 5-hydroxymethyl-uracil (5hmU), 5-hydroxy-5-methylhydantoin (Hyd-Thy), and 5-formyluracil (5fU) have been described as common products with similar yields to cytosine oxidation products¹⁰². Other oxidation products include 5-hydroxy-6-hydrothymine (5-OH-mThy), 5,6-dihydrothymine (5-dihydroT), 5,6-dihydroxycytosine (Cyt-Gly), 5-hydroxy-6-hydrocytosine (5-OHCyt), 5,6-dihydrouracil (H-Ura), alloxan and oxazolone.

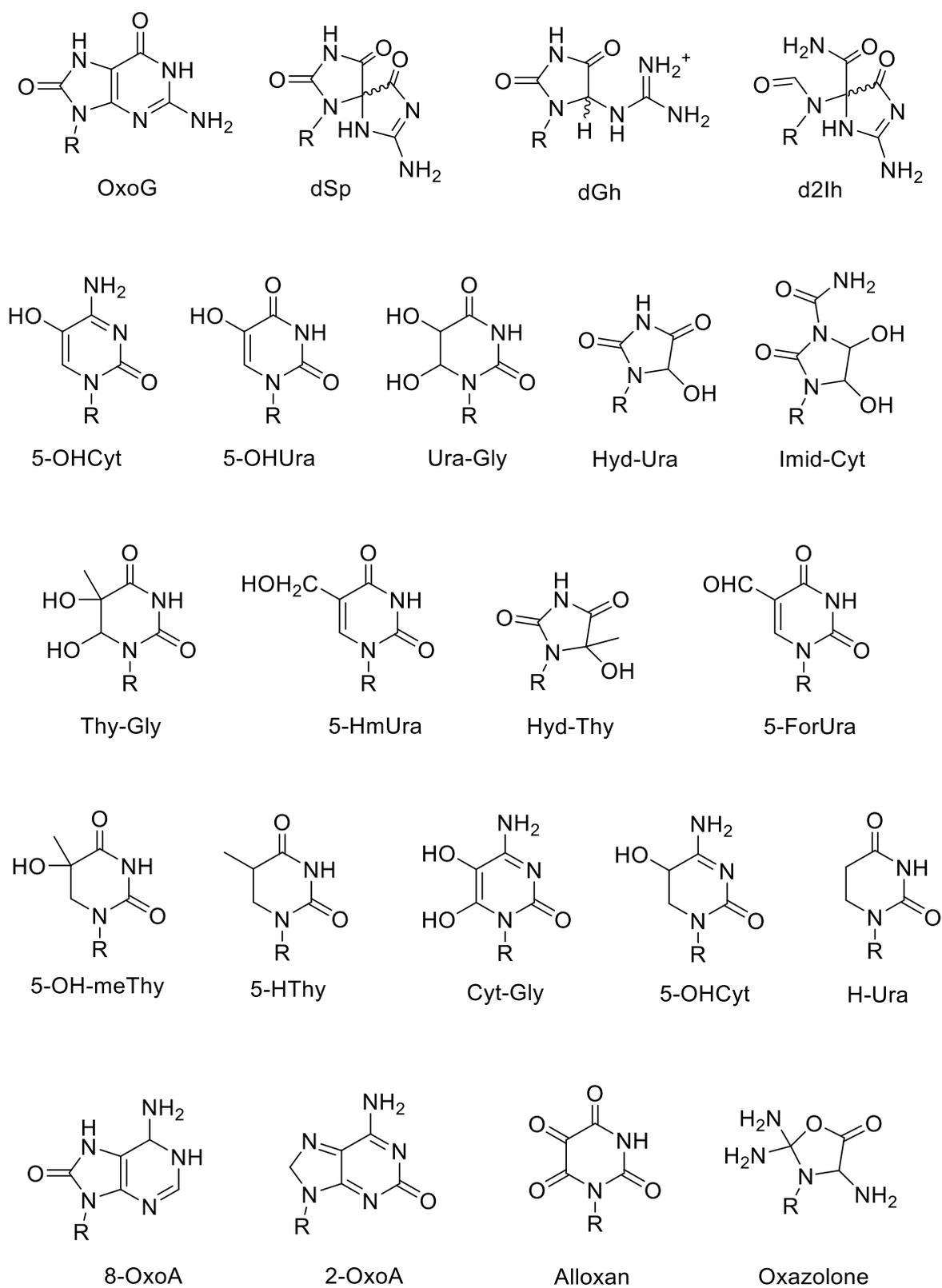


Figure 1.12: Structure diversity of some of the most common damaged bases found in nucleic acids.

1.4. Structural consequences of base modifications

Base modifications can influence the global structure and mechanical properties of the DNA double helix and even higher order-structure in many ways^{103,104}. The presence of new chemical groups for example, which can cause the loss of hydrogen bonds between base pairs or form new ones, can generate possible mispairings¹⁰⁵. Other effects on the structure can as well include changes in the H-bonding patterns of the major groove, modifications of the base pair stacking, changes in the hydration sphere^{106,107} or insertion of steric hindrance¹⁰⁸.

Two techniques which are commonly used to determine nucleic acid structures are X-ray diffraction and nuclear magnetic resonance (NMR) and structures obtained from these techniques are deposited in the Protein Data Bank (PDB). The first reported structure in the PDB that contains a modified base was published in 1982 and contained d(CGCGCG), with all cytosines replaced with the epigenetic marker 5-methylcytosine¹⁰⁶. The structure was later released in 2011 in the PDB. Since then, dozens of structures which investigate the effects of modified bases have been published. Epigenetic markers, for example, have been widely investigated, especially 5-methylcytosine. A total of 51 published structures including one or more 5-methylcytosines have been reported between 1991 and 2021, with 21 more structures containing other epigenetic modifications such as 5-hydroxymethylcytosine or 5-formylcytosine which have also been reported during that time. As for damaged bases, 8-oxoguanine is the most investigated one with a total of 22 published structures reported in the PDB, while a few other damaged bases presented earlier amount for a total of 13 published structures reported. All the structures containing modified bases reported in this work are summarised in Table 1.2, Table 1.3, Table 1.4, and Table 1.5. No structures of protein-nucleic acid complexes have been included.

1.4.1. Effects of epigenetic modifications on the overall structure

As mentioned in the previous paragraph, the oldest structure reported in the PDB in 1982 is the hexamer d(CGCGCG) where all cytosines were replaced by 5-methylcytosines¹⁰⁶. Minor changes in the twist angle between successive base pairs and alteration of some of the interatomic contacts which can be explained by the added hydrophobic contacts of the methyl

groups were reported, although the Z-DNA helix remained similar to the unmethylated d(CGCGCG) structure. Another Z-DNA structure, d((5mC)GTA(5mC)G), this time containing AT base pairs, was reported with similar variations to d((5mC)G(5mC)G(5mC)G)¹⁰⁹. Later, the same sequence with a demethylated thymine was used to investigate stability of Z-DNA, and the presence of a methyl group at the fifth position of the pyrimidine was found to have a destabilising effect¹¹⁰.

Sequences containing only CpG step, or permutations, were studied extensively during the 90s. While a few, as previously seen, formed Z-DNA, many others formed A-DNA structures and this was explained by methylated cytosines stabilising Z-DNA and A-DNA structures over B-DNA. For example, a study published in 1995 showed structures of alternating and non-alternating dG-dC hexanucleotides containing one or multiple methylated cytosines which crystallised as typical A-DNA¹¹¹. Another study published in 1997 did the same but with decamers this time and methylation of cytosines showed typical A-DNA structures¹¹². Finally, another paper published in 1998 pointed out the different hydration sphere between an unmethylated decamer forming A-DNA and a methylated one, but still no major changes compared to a typical A-DNA structure¹⁰⁷.

During the same decade, only two structures reported were B-DNA structures, the decamer d(CCAGGC(5mC)TGG) published in 1992 and the dodecamer d(ACCGC(5mC)GGCGCC)-d(GGCGC(5mC)GGCGG) published in 1997. Both reported that the methylated cytosines did not interfere with helical parameters but changed the base pair locally as well as the hydration sphere^{107,113}. It was not until 2013 that another paper containing a B-DNA structure was published, this time studying the effect of methylated cytosines in the Dickerson-Drew dodecamer (DDD) sequence d(CGCGAATTCGCG)¹¹⁴.

Table 1.2: Summary of all structures containing the base 5-methylcytosine that have been published on the

PDB.

PDB entry	Technique	Resolution (Å)	Structure type	Reference	PDB entry	Technique	Resolution (Å)	Structure type	Reference
2D25	X-ray	1.75	B-DNA	113	1IH4	X-ray	1.90	B-A DNA	115
1D41	X-ray	1.30	Z-DNA	110	1IH3	X-ray	2.40	B-A DNA	115
1D40	X-ray	1.30	Z-DNA	110	1L6B	X-ray	1.50	Holliday Junction	116
229D	NMR	-	-	124	1R41	X-ray	1.90	A-DNA	117
213D	X-ray	1.60	A-DNA	125	1R3Z	X-ray	1.70	A-DNA	117
254D	X-Ray	1.90	A-DNA	111	1S1L	X-ray	2.20	Holliday Junction	118
256D	X-ray	2.20	A-DNA	111	1VTV	X-ray	1.30	Z-DNA	106
275D	X-ray	2.00	A-DNA	111	1VTW	X-ray	1.20	Z-DNA	109
297D	X-ray	2.50	B-DNA	126	4GJU	X-ray	1.41	B-DNA	114
329D	X-ray	2.70	B-DNA	127	4GLG	X-ray	1.72	B-DNA	114
285D	X-ray	2.50	B-DNA	126	4C63	X-ray	1.32	B-DNA	119
324D	X-ray	2.15	A-DNA	112	4MKW	X-ray	1.22	B-DNA	120
325D	X-ray	2.50	A-DNA	126	5DSA	X-ray	1.69	Holliday Junction	121
326D	X-ray	2.15	A-DNA	126	5L06	NMR	-	B-DNA	122
327D	X-ray	1.94	A-DNA	128	5L2G	NMR	-	B-DNA	122
322D	X-ray	2.50	A-DNA	126	5UZ2	NMR	-	B-DNA	122
323D	X-ray	2.15	A-DNA	126	5UZ3	NMR	-	B-DNA	122
313D	X-ray	1.68	Z-DNA	129	5UZ1	NMR	-	B-DNA	122
384D	X-ray	2.15	A-DNA	107	5TRN	NMR	-	B-DNA	122
383D	X-ray	1.70	A-DNA	107	6ALS	NMR	-	B-DNA	122
400D	X-ray	1.65	Z-DNA	130	6ALU	NMR	-	B-DNA	122
1EL2	NMR	-	-	131	6ALT	NMR	-	B-DNA	122
1ELN	NMR	-	-	131	6JV5	X-ray	1.40	B-DNA	65
1F6E	X-ray	2.00	A-DNA	132	7D0Y	NMR	-	-	123
1F6I	X-ray	2.20	E-DNA	132	7D0X	NMR	-	-	123
1FV7	NMR	-	B-Z junction	133					

Hybrid structures formed by the presence of 5-methylcytosine have also been reported. For example, a paper describing a right-handed Z-DNA helix embedded in a B-DNA matrix was published in 2000¹³³. A year later, a study investigated the transition from a B-DNA to a A-DNA structure using methylated and unmethylated and brominated CpG sequences¹¹⁵. Holliday junctions were also reported three times in 2002, 2004 and 2016 and the presence of the methyl group significantly affects the local structure and disrupts hydrogen bonding at the junction^{116,118,121}. Recently, the effect of adding oxidised bases to methylated sequences has also been considered for structural studies of CpG sequences. Oxidation was reported to increase base pair breathing, to induce transitions in the backbone 3' to the oxoG and to reduce variability of shift and tilt helical parameters which was suggested to modify interactions with enzymes specifically recognizing methylated CpG dinucleotides¹²².

In the 2010s, the revived interest in 5-hydroxymethylcytosine led teams to crystallise sequences that included this base to see its effect on the overall structure. At the time, it was proposed that while 5-methylcytosine stabilised the duplex, 5-hydroxymethylcytosine reversed this stabilisation to give a structure very similar to the unmodified form¹³⁴. Crystallisation of the DDD at different positions of the strand with 5-methylcytosine and 5-hydroxymethylcytosine in 2013 showed no influence on the structure, as well as no change in the overall thermodynamics, thus the authors concluded that the 5 position of cytosine is an ideal place to encode epigenetic information¹¹⁴. In the following year, the structure of the same sequence with the same 5-hydroxymethylcytosine modification was solved at higher resolution^{119,135}. The same conclusion was reached, 5-hydroxymethylcytosine has little effect on the global and local geometry of the duplex. 5-hydroxymethylcytosine was then incorporated into a Holliday junction forming sequence to investigate its impact on the stability considering the modified base's role in promoting recombination. The authors concluded that there was a minimal effect on the overall thermodynamic stability of the junction, with a junction core even more accessible than in the 5-methylcytosine modified structure¹²¹ (Figure 1.13).

Table 1.3: Summary of all structures containing the base 5-hydroxymethylcytosine (top) and 5-formylcytosine (bottom) that have been published on the PDB.

Damaged base	PDB entry	Technique	Resolution (Å)	Structure type	Reference
5hmC	4GLC	X-ray	1.83	B-DNA	114
	4GLH	X-ray	1.66	B-DNA	114
	4HLI	X-ray	1.99	B-DNA	114
	4C5X	X-ray	1.30	B-DNA	119
	4I9V	X-ray	1.02	B-DNA	135
	5DSB	X-ray	1.50	Holliday Junction	121
	6JV3	X-ray	2.85	Holliday Junction	65
5fC	4QKK	X-ray	1.40	A-DNA	136
	4QC7	X-ray	1.90	B-DNA	135
	5HN2	X-ray	1.50	RNA	137
	5HNQ	X-ray	2.40	RNA	137
	5HNJ	X-ray	1.24	RNA	137
	5MVU	X-ray	2.30	A-DNA	138
	5ZAS	X-ray	1.56	A-DNA	65

Table 1.4: Summary of all structures containing the base 5-carboxylcytosine (top) and 6-methyladenine (bottom) that have been published in the PDB.

Damaged base	PDB entry	Technique	Resolution (Å)	Structure type	Reference
5caC	4PWM	X-ray	1.95	B-DNA	135
	5ZAT	X-ray	1.06	B-DNA	65
6mA	4DNB	X-ray	2.00	B-DNA	139
	1DA3	X-ray	2.00	B-DNA	140
	1OQ2	NMR		B-DNA	141
	1UAB	NMR		B-DNA	141
	2KAL	NMR		B-DNA	142

5-Formylcytosine also had its effects on the structural stability of DNA investigated. The first crystallographic structure was published in 2015 and the authors claimed that this base was the first to significantly change the B-DNA, naming this unusual structure F-DNA¹³⁶. However, these findings were challenged a few years later when a group of researchers concluded using NMR and X-ray crystallography that it, in fact, did not change the global structure of the DNA¹³⁸. Furthermore, in 2015, another team also crystallised the DDD that

included 5fC, not reporting any unusual structure¹³⁵ (Figure 1.13). In 2019 however, a structure containing 5fC was reported and the authors concluded that its incorporation in the sequence induced the geometry alteration of the DNA minor groove, alteration that could be recognised distinctively by the enzyme thymine DNA glycosylase⁶⁵. The same conclusion was reached for 5-carboxylcytosine, which is the least reported 5-methylcytosine oxidation product on the PDB among the four bases which have been characterised in the methylation-demethylation pathway (Figure 1.13).

In contrast, only a few DNA structures containing 6-methyladenine have been reported on the PDB, compared to the number of structures which involve cytosine-modified bases. First thought to be mainly present in prokaryotic species after being discovered in *E. coli*¹⁴³, the structural studies done at the time were to understand the effects of the methylation as it prevented the cleavage of DNA by restriction endonucleases. The first result showed no alteration that could explain how the restriction enzyme EcoRI recognises the modification¹³⁹. It is the specific recognition of the hemimethylated GATC site by the SeqA protein that led researchers to look at the structure and dynamics of hemimethylated sequences. The authors reported that even though the helix was still not altered significantly, NMR data showed the local site to slowly interconvert between *trans* and *cis*¹⁴¹. Furthermore, the team described a narrow major groove, form that could induce a specific recognition of the site by the SeqA protein. A second team reported similar results with an hemimethylated duplex containing a unique groove structure similar to the one reported previously and observed its disappearance in fully methylated duplexes¹⁴². No structure including this modification has been reported in the PDB since 2008, but recent reports of 6-methyladenine in eukaryotic organisms, especially in neuronal tissues and during early embryo development could rekindle interest in DNA structures post-methylation.

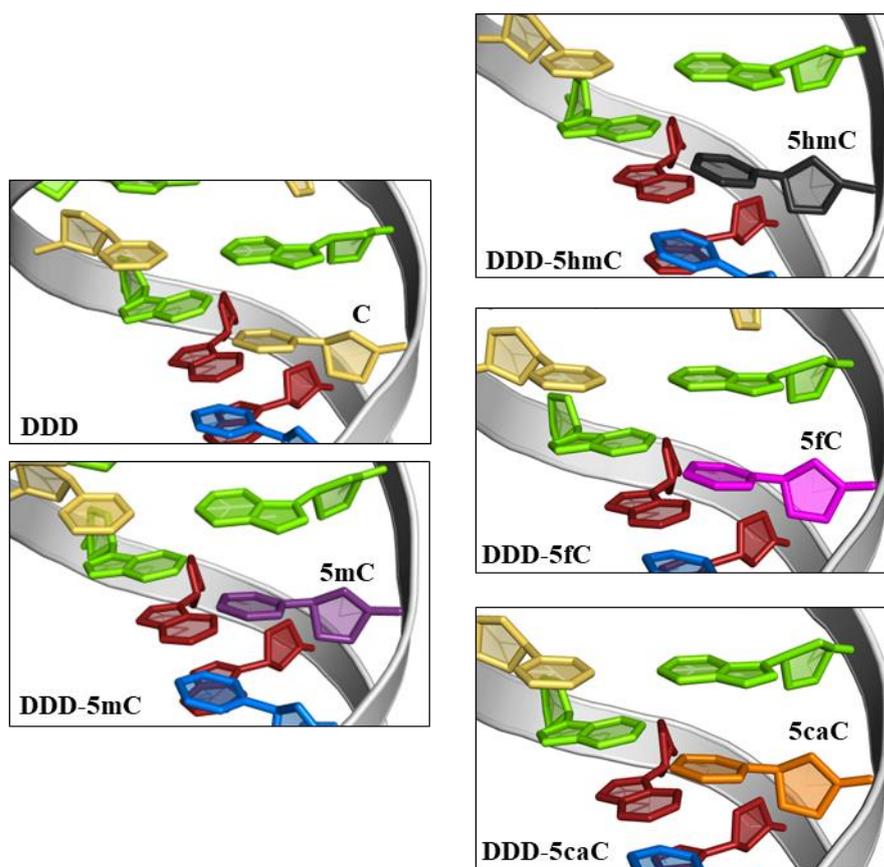


Figure 1.13: Close-up of the of the crystal structures of the Dickerson-Drew dodecamers with different bases from the methylation-demethylation pathway. The backbone chain is in grey. Nucleotides are coloured following convention, guanine bases in green, adenine bases in red and thymine bases in blue. The modified nucleotides are coloured as follow: 5-methylcytosine in purple, 5-hydroxymethylcytosine in black, 5-formylcytosine in pink and 5-carboxylcytosine in orange. Only local parameters have been modified by the presence of modified bases. The structures shown are taken from entry 7BNA¹⁴⁴ (DDD), 4GLG¹¹⁴ (DDD-5C), 4I9V¹³⁵ (DDD-5hmC), 4QC7¹³⁵ (DDD-5fC) and 4PWM¹³⁵ (DDD-5caC) from the Protein Data Bank³⁰ (PDB).

1.4.2. Effects of damaged bases on the overall structure of nucleic acids

Alongside structures including epigenetic bases to investigate their effects on overall and local nucleic acid structures, effects of lesions have also been examined and reported on the PDB. The guanine oxidation product 8-oxoguanine, oxoG, is the most commonly used oxidative stress biomarker and its impact on DNA conformation has been studied since the mid-90's.

To prevent transversions from one base to another during replication, repair enzymes are dedicated to correct mispairing. The MutM protein found in *E. coli* for example will remove the oxoG lesions from the strand in oxoG:C base pairs¹⁴⁵. Of course, recognition of damage by the repair enzyme is important and crystallisation of a duplex containing that base pair gave insight to a unique pattern of hydrogen donors and acceptors within the major groove which could explain how MutM can recognize it. The duplex conformation did not sustain any major change otherwise¹⁴⁶. Similarly, repair enzymes found in human cells that remove oxoG were also found to remove the damaged base from oxoG:C, oxoG:T and oxoG:G base pairs, but weren't as efficient at removing them in oxoG:A base pairs¹⁴⁷. Structural data showed that the oxoG:A base pair was structurally similar to a A:T base pair, rendering them nearly indistinguishable from one another¹⁴⁷ (Figure 1.14). Another similar study also investigated the structural effect of oxoG when paired with another guanine and concluded that the oxoG:G base pair formed a Hoogsteen base-pair, but the duplex remained structurally similar to the one that is undamaged¹⁴⁸. Subsequently, the interest shifted to the effects of abasic sites close to oxoG that could prevent the oxoG base excision by the Fpg protein. However, structural information only couldn't explain why these damage sites could be recognized by the enzyme^{149,150}. Similarly, another team inserted an oxoG into the Dickerson-Drew dodecamer (DDD) to gain insight as to why the EcoRI restriction enzyme's activity was decreased in the presence of this base, as the DDD contains the enzyme's recognition site, GAATTC. The authors observed substantial alterations in the sugar-phosphate backbone conformation and correlated the enzymatic disruption with the DNA backbone shift¹⁵¹.

The effects of oxoguanine bases on higher-order structures such as G-quadruplexes has also been investigated in the past decade¹⁰⁴. This comes from the suggestion that G-rich tracts act as oxidative sinkholes, protecting the rest of the genome¹⁵². OxoG is believed to be detrimental to this DNA structure as the loss of hydrogen bonds can destabilise the whole structure. A first study on the human telomere sequence (hTel) showed that oxoG substitutions caused either minor structural adjustment or major conformational shift¹⁰⁴. However, when the same team looked at a quadruplex forming sequence present upstream to the P1 promoter of the BCL2 gene and substituted some of the guanines by oxoG, they discovered that the damaged base had a stabilising effect on the overall structure. A potential novel regulatory role of oxidative stress in general was suggested as a conclusion¹⁵³. Finally, another G quadruplex structure containing

oxoG was also published, this time paired with another damaged base, xanthine. Reestablishment of the hydrogen bond between complimentary bases was observed while each damaged base on its own destabilised the whole structure¹⁵⁴.

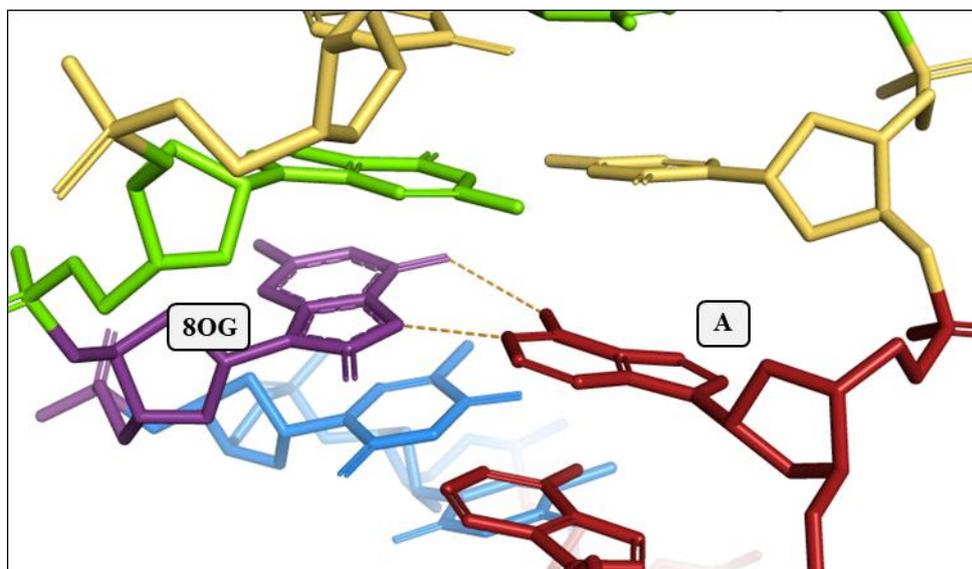


Figure 1.14: Close-up of the of the crystal structures of a dodecamer with an A:oxoG base pair. Nucleotides are coloured following convention, guanine bases in green, adenine bases in red and thymine bases in blue. The 8-oxoguanine base is in purple. Only local parameters have been modified by the presence of modified bases. The structure shown is taken from entry 178D¹⁴⁷ from the Protein Data Bank³⁰ (PDB).

Structures that include other damaged bases than oxoG have also been reported. Oxoadenine has been studied similarly to oxoG with two structures including the oxoA:G and oxoA:T base pairs which induced local changes in the duplex^{155,156}. Furthermore, the lesion 8-methylguanine has been incorporated in a Z-DNA forming sequence and a similar stabilisation to 5-methylcytosine has been observed¹⁵⁷. Finally, the effects of the 6-carboxymethylguanine alkylation damage have also been described on two separate occasions in 2013 and 2014, both considering the presence of this damage in colorectal DNA which could be linked to colorectal cancer^{158,159}.

Table 1.5: Summary of all structures containing the base 8-oxoguanine that have been published on the PDB.

Damaged base	PDB entry	Technique	Resolution (Å)	Structure type	Reference
oxoG	178D	X-ray	2.50	B-DNA	147
	183D	X-ray	1.60	B-DNA	146
	1N2W	NMR		B-DNA	148
	2M40	NMR		B-DNA	150
	2M44	NMR		B-DNA	150
	2M43	NMR		B-DNA	150
	2M3P	NMR		B-DNA	150
	2M3Y	NMR		B-DNA	150
	2MWZ	NMR		G-Quad	154
	5IV1	NMR		B-DNA	151
	5IZP	NMR		B-DNA	151
	5HQQ	NMR		B-DNA	149
	5HQF	NMR		B-DNA	149
	5UZ2	NMR		B-DNA	122
	5UZ3	NMR		B-DNA	122
	5UZ1	NMR		B-DNA	122
	5TRN	NMR		B-DNA	122
	6ALS	NMR		B-DNA	122
	6ALU	NMR		B-DNA	122
	6IA0	NMR		G-Quad	104
	6IA4	NMR		G-Quad	104
	6ZX6	NMR		G-Quad	153

Table 1.6: Summary of all structures containing the bases 8-oxoadenine, 8-methylguanine and 6-carboxymethylguanine that have been published on the PDB.

Damaged base	PDB entry	Technique	Resolution (Å)	Structure type	Reference
oxoA	1D75	X-ray	2.80	B-DNA	155
	1FJB	NMR		B-DNA	156
8mG	1TNE	NMR		Z-DNA	157
6CaG	4IJ0	X-ray	1.54	B-DNA	158
	4ITD	X-ray	1.94	B-DNA	158
	4O5W	X-ray	1.60	B-DNA	159
	4O5X	X-ray	1.60	B-DNA	159
	4O5Y	X-ray	1.75	B-DNA	159
	4O5Z	X-ray	1.75	B-DNA	159

1.5. Incorporation of transition metals into nucleic acid structures: how the identification of transition metal centres can help identify potential damage centres

The reaction between transition metals, especially copper and iron, and hydrogen peroxide, or Fenton reaction, is an important factor contributing to nucleic acids damage *in vitro* and *in vivo*¹⁶⁰⁻¹⁶². Furthermore, since transition metals bind to nucleic acids through the phosphate backbone and endocyclic atoms *e.g.* at the N7 atoms of purine bases¹⁶³, it is possible hydroxyl radicals HO•, a short-lived oxygen species produced by the Fenton reaction, reacts at diffusion rates with nearby bases¹⁶². In summary, transition-metals' binding preferences with nucleic acids could be correlated to damage centres. This observation highlights how nucleic acid structures containing transition-metal binding sites are essential, particularly to understand where damage might happen in the overall structure at an atomic scale.

Many nucleic acid structures containing free transition-metal ions have been reported in the PDB in the past decades. Some of the most represented transition metals include Mn²⁺, Zn²⁺ and Co²⁺ with 61, 34 and 31 published structures respectively. On the contrary, other transition metals such as Ni²⁺ and Cu²⁺ are underrepresented with only 7 and 9 structures published respectively. In the interest of focusing on nucleic acids - transition metal interactions, no structures containing proteins have been included.

Few nucleic acid structures reported in the PDB contain nickel (II) ions, but the few that do give an insight in how divalent transition metals interact with nucleic acids. As expected, in the majority of the structures reported, Ni²⁺ ions would bind preferentially to the N7 position of guanines¹⁶⁴⁻¹⁷⁰. However, it is worth noting that phosphate-Ni²⁺-guanine¹⁶⁸ and guanine-Ni²⁺-neighbour duplex^{164,165} bridges are common occurrences in reported X-ray structures due to crystal packing (Figure 1.15).

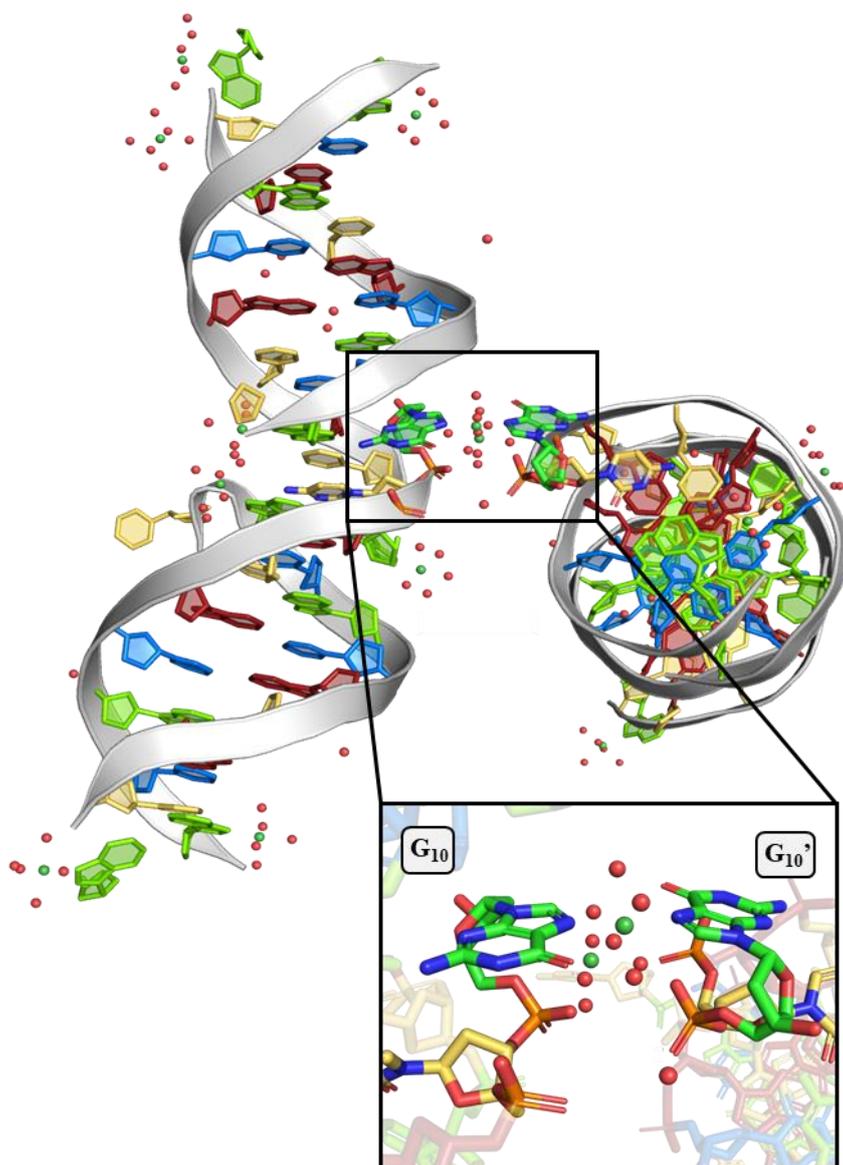


Figure 1.15: Guanine- Ni^{2+} -guanine bridges were reported in the crystal structure of the B-DNA duplex (PDB ID: 1G3V¹⁶⁶) between symmetry-related duplexes. The backbone chain is in light grey. Nucleotides are coloured following convention, guanine bases in green, adenine bases in red and thymine bases in blue

Co^{2+} ions and Zn^{2+} ions also present similar proprieties when introduced in crystal systems, with terminal bases sometimes flipping out of the way as a result¹⁷¹⁻¹⁷⁴. These can occur among symmetry-related bases in crystal structures when the transition metal is directly present in the crystallisation conditions, as is often the case in commercially available screens. The presence of transition metals in screens is generally due to their practical use in X-ray crystallography. They can be utilised as heavy atoms to solve structures with the Single and Multi-wavelength

Anomalous Dispersion (SAD/MAD) method^{175,176}, a method applied to solve the structure by calculating their position in the structure^{177–180}. However, while many of these studies do not focus primarily on nucleic acids - transition metal interactions^{179–181}, it was noted that the binding can influence crystallisation¹⁸² and stabilise the structures observed¹⁸¹.

Transition-metals availability to hydrogen peroxide is as important as transition-metals binding preferences in context with the Fenton reaction as a source of nucleic-acid damage, especially if the presence of transition metals influence structures as observed in nucleic-acid crystal systems. This is not necessarily a problem with proteins, and in the past, hydrogen peroxide has been introduced to crystal systems to investigate biological processes such as catalysis in metalloprotein containing transition metals in their active sites^{183–187}. However, the presence of base – transition metal bridges in nucleic acid systems might influence availability of the transition metal to hydrogen peroxide and prevent the reaction. Fenton-based oxidative damage in proteins has also been investigated using crystal systems¹⁸⁸, a method which has surprisingly never been applied to nucleic-acid crystal systems.

1.6. Aims

Structural studies resulting from techniques such as X-ray crystallography and NMR are powerful tools to understand processes which rely on pattern recognition. Almost all the studies discussed here were to identify structural changes after modifications and the effects on enzymes recognition, as enzymes bind to the substrate through its active site, and small changes in the groove structure at damage sites were suggested to be what is recognised by enzymes. However, it is not known if DNA modification when damage occurs by chemical species, occurs randomly on the DNA strand, or if there could be hotspots where modifications would ensue. For this reason, the aim of this project is to understand the role nucleic acid structure plays in directing the effect of damaging chemical processes and will explore the susceptibility of different structures to damage. Furthermore, this project will focus particularly on Fenton-mediated nucleic acid damage, a reaction catalysed by transition metals. For this, both X-ray crystallography and circular dichroism (CD) spectroscopy are appropriate techniques as they allow the studying of interactions between transition metals and nucleic acid. X-ray

crystallography and CD also offer environments compatible with the introduction of hydrogen peroxide to study the structural effects of damage on nucleic acids in the presence of transition metals - i.e., crystalline environment and aqueous solution environment. Using NMR was not considered as getting a structure with transition metals using this method can be challenging due to the magnetic properties of some of them¹⁸⁹.

As discussed earlier in the previous paragraph, transition metals have been known to catalyse the Fenton reaction, and previous work has shown preferential binding of metal ions to different bases. This is however unknown if catalysis occurring at these spots would damage the closest base or if the bound transition metal would protect the base it is bound to form damage. Crystallisation of a variety of structures with these transition metals may allow for the identification of damage hotspots. To enable this, new methodology for inducing in-crystal reactions needs to be developed, and crystal lifetimes in reaction mixtures be established.

Results coming from structural studies could contribute to a better understanding of how nucleic acids can be modified and which structures are more likely to be damaged. Additionally, with the knowledge of nucleic acid sequences, it could be possible to predict parts that are potential sensitive spots in longer strands. Furthermore, studying transition metals involved in oxidative damage and oxidative stress may lead to a better understanding of what could occur to nucleic acids during the breakdown of metalloregulation processes in neurodegenerative diseases such as Alzheimer's.

Specific aims of this work are to:

1. Show which method is required to obtain ordered metal sites in crystals with minimum structural perturbation.

The addition of transition metal ions to crystal systems containing biomolecules can generally be achieved in two ways: Crystallisation from a solution containing both the metal ion and the biomolecules, a method referred to as cocrystallisation, and soaking a grown crystal with a solution containing the metal ion of interest. The metal sites availability to reactants from the solvent channel and little structural perturbations would be considered crucial to investigate oxidative damage.

2. Investigate the binding preferences of transition metals capable of catalysing the Fenton reaction

Copper and iron have been previously investigated for their ability to damage DNA by reacting with hydrogen peroxide, but little is known about the preferred binding sites of these two transition metals.

3. Investigate the effects of soaking a crystal containing ordered metal centres in hydrogen peroxide

Previous studies have described protein crystal systems which were soaked in hydrogen peroxide, but this has never been attempted with a DNA crystal. Soaking a DNA crystal in hydrogen peroxide could help shed light on to where DNA damage occur.

4. Show the sensitivity of specific structures to oxidative damage

G-quadruplexes are particularly sensitive to oxidative damage and circular dichroism can be used to monitor changes in G-quadruplexes topologies, as one of the consequences of guanine oxidation is the protonation at the N7 position of guanines, a modification which can prevent Hoogsteen pairings between the oxidised guanine and the other guanines in the tetrad and destabilise the whole structure

1.7. References

1. Dahm, R. Friedrich Miescher and the discovery of DNA. *Dev. Biol.* **278**, 274–288 (2005).
2. Judd, B. H. Nucleic Acids as Genetic Material. in *eLS* (2001).
3. Sinden, R. R. Chapter 1 - Introduction to the Structure, Properties, and Reactions of DNA. in *DNA Structure and Function* 1–57 (Academic Press, San Diego, 1994).
4. Minchin, S. & Lodge, J. Understanding biochemistry: structure and function of nucleic acids. *Essays Biochem.* **63**, 433–456 (2019).
5. Travers, A. & Muskhelishvili, G. DNA structure and function. *FEBS J.* **282**, 2279–2295 (2015).
6. Marin-Gonzalez, A., Vilhena, J. G., Perez, R. & Moreno-Herrero, F. A molecular view of DNA flexibility. *Q. Rev. Biophys.* **54**, e8 (2021).
7. Thompson, J. M. T. & Travers, A. A. The structural basis of DNA flexibility. *Philos. Trans. R. Soc. Lond. Ser. Math. Phys. Eng. Sci.* **362**, 1423–1438 (2004).
8. Hakem, R. DNA-damage repair; the good, the bad, and the ugly. *EMBO J.* **27**, 589–605 (2008).
9. Vella, F. Nucleic Acids in Chemistry and Biology. *Biochem. Educ.* **19**, 97–98 (1991).
10. Lenz, S. A. P., Kellie, J. L. & Wetmore, S. D. Glycosidic Bond Cleavage in DNA Nucleosides: Effect of Nucleobase Damage and Activation on the Mechanism and Barrier. *J. Phys. Chem. B* **119**, 15601–15612 (2015).
11. Huang, M., Giese, T. J., Lee, T.-S. & York, D. M. Improvement of DNA and RNA Sugar Pucker Profiles from Semiempirical Quantum Methods. *J. Chem. Theory Comput.* **10**, 1538–1545 (2014).
12. Altona, C. & Sundaralingam, M. Conformational analysis of the sugar ring in nucleosides and nucleotides. New description using the concept of pseudorotation. *J. Am. Chem. Soc.* **94**, 8205–8212 (1972).

13. Chargaff, E. Chemical specificity of nucleic acids and mechanism of their enzymatic degradation. *Experientia* **6**, 201–209 (1950).
14. Watson, J. D. & Crick, F. H. C. Molecular Structure of Nucleic Acids: A Structure for Deoxyribose Nucleic Acid. *Nature* **171**, 737–738 (1953).
15. Hoogsteen, K. The crystal and molecular structure of a hydrogen-bonded complex between 1-methylthymine and 9-methyladenine. *Acta Crystallogr.* **16**, 907–916 (1963).
16. Crick, F. H. C. Codon—anticodon pairing: The wobble hypothesis. *J. Mol. Biol.* **19**, 548–555 (1966).
17. Frank-Kamenetskii, M. D. & Mirkin, S. M. Triplex DNA Structures. *Annu. Rev. Biochem.* **64**, 65–95 (1995).
18. Burge, S., Parkinson, G. N., Hazel, P., Todd, A. K. & Neidle, S. Quadruplex DNA: sequence, topology and structure. *Nucleic Acids Res.* **34**, 5402–5415 (2006).
19. Varani, G. & McClain, W. H. The G·U wobble base pair. *EMBO Rep.* **1**, 18–23 (2000).
20. Leontis, N. B., Stombaugh, J. & Westhof, E. The non-Watson–Crick base pairs and their associated isostericity matrices. *Nucleic Acids Res.* **30**, 3497–3531 (2002).
21. Lu, X. & Olson, W. K. 3DNA: a software package for the analysis, rebuilding and visualization of three-dimensional nucleic acid structures. *Nucleic Acids Res.* **31**, 5108–5121 (2003).
22. Franklin, R. E. & Gosling, R. G. The structure of sodium thymonucleate fibres. I. The influence of water content. *Acta Crystallogr.* **6**, 673–677 (1953).
23. Lucas, A. A. A-DNA and B-DNA: Comparing Their Historical X-ray Fiber Diffraction Images. *J. Chem. Educ.* **85**, 737 (2008).
24. Wang, A. H.-J. *et al.* Molecular structure of a left-handed double helical DNA fragment at atomic resolution. *Nature* **282**, 680–686 (1979).
25. Pastor, N. The B- to A-DNA Transition and the Reorganization of Solvent at the DNA Surface. *Biophys. J.* **88**, 3262–3275 (2005).
26. Egli, M. Chapter 2: DNA and RNA Structure. in *Nucleic Acids in Chemistry and Biology* 20–95 (2022).

27. Bingman, C., Jain, S., Zon, G. & Sundaralingam, M. Crystal and molecular structure of the alternating dodecamer d(GCGTACGTACGC) in the A-DNA form: comparison with the isomorphous non-alternating dodecamer d(CCGTACGTACGG). *Nucleic Acids Res.* **20**, 6637–6647 (1992).
28. Drew, H. R. *et al.* Structure of a B-DNA dodecamer: conformation and dynamics. *Proc. Natl. Acad. Sci.* **78**, 2179–2183 (1981).
29. Brzezinski, K. *et al.* High regularity of Z-DNA revealed by ultra high-resolution crystal structure at 0.55 Å†. *Nucleic Acids Res.* **39**, 6238–6248 (2011).
30. Bernstein, F. C. *et al.* The protein data bank: A computer-based archival file for macromolecular structures. *J. Mol. Biol.* **112**, 535–542 (1977).
31. Vasquez, K. M. & Glazer, P. M. Triplex-forming oligonucleotides: principles and applications. *Q. Rev. Biophys.* **35**, 89–107 (2002).
32. Bevilacqua, P. C. & Blose, J. M. Structures, Kinetics, Thermodynamics, and Biological Functions of RNA Hairpins. *Annu. Rev. Phys. Chem.* **59**, 79–103 (2008).
33. Bikard, D., Loot, C., Baharoglu, Z. & Mazel, D. Folded DNA in Action: Hairpin Formation and Biological Functions in Prokaryotes. *Microbiol. Mol. Biol. Rev.* **74**, 570–588 (2010).
34. Watson, J., Hays, F. A. & Ho, P. S. Definitions and analysis of DNA Holliday junction geometry. *Nucleic Acids Res.* **32**, 3017–3027 (2004).
35. Postow, L. *et al.* Positive Torsional Strain Causes the Formation of a Four-way Junction at Replication Forks. *J. Biol. Chem.* **276**, 2790–2796 (2001).
36. Spiegel, J., Adhikari, S. & Balasubramanian, S. The Structure and Function of DNA G-Quadruplexes. *Trends Chem.* **2**, 123–136 (2020).
37. Zhao, J., Bacolla, A., Wang, G. & Vasquez, K. M. Non-B DNA structure-induced genetic instability and evolution. *Cell. Mol. Life Sci.* **67**, 43–62 (2010).
38. Wang, G. & Vasquez, K. M. Naturally occurring H-DNA-forming sequences are mutagenic in mammalian cells. *Proc. Natl. Acad. Sci.* **101**, 13448–13453 (2004).

39. Grigg, J. C., Shumayrikh, N. & Sen, D. G-Quadruplex Structures Formed by Expanded Hexanucleotide Repeat RNA and DNA from the Neurodegenerative Disease-Linked C9orf72 Gene Efficiently Sequester and Activate Heme. *PLOS ONE* **9**, e106449 (2014).
40. Mergny, J.-L., Lacroix, L., Han, X., Leroy, J.-L. & Helene, C. Intramolecular Folding of Pyrimidine Oligodeoxynucleotides into an i-DNA Motif. *J. Am. Chem. Soc.* **117**, 8887–8898 (1995).
41. Rhodes, D. & Lipps, H. J. G-quadruplexes and their regulatory roles in biology. *Nucleic Acids Res.* **43**, 8627–8637 (2015).
42. Cheng, M. *et al.* Loop permutation affects the topology and stability of G-quadruplexes. *Nucleic Acids Res.* **46**, 9264–9275 (2018).
43. Parkinson, G. N., Lee, M. P. H. & Neidle, S. Crystal structure of parallel quadruplexes from human telomeric DNA. *Nature* **417**, 876–880 (2002).
44. Zeraati, M. *et al.* I-motif DNA structures are formed in the nuclei of human cells. *Nat. Chem.* **10**, 631–637 (2018).
45. Weil, J. *et al.* Stabilization of the i-motif by intramolecular adenine–adenine–thymine base triple in the β -structure of d(ACCCCT). *Acta Crystallogr. D Biol. Crystallogr.* **55**, 422–429 (1999).
46. Day, H. A., Pavlou, P. & Waller, Z. A. E. i-Motif DNA: Structure, stability and targeting with ligands. *Bioorg. Med. Chem.* **22**, 4407–4418 (2014).
47. Kumar, S., Chinnusamy, V. & Mohapatra, T. Epigenetics of Modified DNA Bases: 5-Methylcytosine and Beyond. *Front. Genet.* **9**, (2018).
48. Lindahl, T. Instability and decay of the primary structure of DNA. *Nature* **362**, 709–715 (1993).
49. Martin, L. J. DNA Damage and Repair: Relevance to Mechanisms of Neurodegeneration. *J. Neuropathol. Exp. Neurol.* **67**, 377–387 (2008).
50. Portela, A. & Esteller, M. Epigenetic modifications and human disease. *Nat. Biotechnol.* **28**, 1057–1068 (2010).
51. Waddington, C. H. The Epigenotype. *Int. J. Epidemiol.* **41**, 10–13 (2012).

52. Waddington, C. H. Towards a Theoretical Biology. *Nature* **218**, 525–527 (1968).
53. Breiling, A. & Lyko, F. Epigenetic regulatory functions of DNA modifications: 5-methylcytosine and beyond. *Epigenetics Chromatin* **8**, 24 (2015).
54. Field, S. F. *et al.* Accurate Measurement of 5-Methylcytosine and 5-Hydroxymethylcytosine in Human Cerebellum DNA by Oxidative Bisulfite on an Array (OxBS-Array). *PLOS ONE* **10**, e0118202 (2015).
55. Bird, A., Taggart, M., Frommer, M., Miller, O. J. & Macleod, D. A fraction of the mouse genome that is derived from islands of nonmethylated, CpG-rich DNA. *Cell* **40**, 91–99 (1985).
56. Cooper, D. N., Taggart, M. H. & Bird, A. P. Unmethlated domains in vertebrate DNA. *Nucleic Acids Res.* **11**, 647–658 (1983).
57. Lister, R. *et al.* Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature* **462**, 315–322 (2009).
58. Tahiliani, M. *et al.* Conversion of 5-Methylcytosine to 5-Hydroxymethylcytosine in Mammalian DNA by MLL Partner TET1. *Science* **324**, 930–935 (2009).
59. Ito, S. *et al.* Tet Proteins Can Convert 5-Methylcytosine to 5-Formylcytosine and 5-Carboxylcytosine. *Science* **333**, 1300–1303 (2011).
60. Marinus, M. G. & Morris, N. R. Biological function for 6-methyladenine residues in the DNA of *Escherichia coli* K12. *J. Mol. Biol.* **85**, 309–322 (1974).
61. Luo, G.-Z., Blanco, M. A., Greer, E. L., He, C. & Shi, Y. DNA N6-methyladenine: a new epigenetic mark in eukaryotes? *Nat. Rev. Mol. Cell Biol.* **16**, 705–710 (2015).
62. Hermann, A., Schmitt, S. & Jeltsch, A. The Human Dnmt2 Has Residual DNA-(Cytosine-C5) Methyltransferase Activity. *J. Biol. Chem.* **278**, 31717–31721 (2003).
63. Bokar, J. A., Shambaugh, M. E., Polayes, D., Matera, A. G. & Rottman, F. M. Purification and cDNA cloning of the AdoMet-binding subunit of the human mRNA (N6-adenosine)-methyltransferase. *RNA* **3**, 1233–1247 (1997).
64. Tang, L.-Y. *et al.* The Eukaryotic DNMT2 Genes Encode a New Class of Cytosine-5 DNA Methyltransferases. *J. Biol. Chem.* **278**, 33613–33616 (2003).

65. Fu, T. *et al.* Thymine DNA glycosylase recognizes the geometry alteration of minor grooves induced by 5-formylcytosine and 5-carboxylcytosine. *Chem. Sci.* **10**, 7407–7417 (2019).
66. Ba, X. & Boldogh, I. 8-Oxoguanine DNA glycosylase 1: Beyond repair of the oxidatively modified base lesions. *Redox Biol.* **14**, 669–678 (2018).
67. Zhang, G. *et al.* N6-Methyladenine DNA Modification in *Drosophila*. *Cell* **161**, 893–906 (2015).
68. Greer, E. L. *et al.* DNA Methylation on N6-Adenine in *C. elegans*. *Cell* **161**, 868–878 (2015).
69. Zhang, M. *et al.* Mammalian ALKBH1 serves as an N6-mA demethylase of unpairing DNA. *Cell Res.* **30**, 197–210 (2020).
70. Povirk, L. F. & Shuker, D. E. DNA damage and mutagenesis induced by nitrogen mustards. *Mutat. Res. Genet. Toxicol.* **318**, 205–226 (1994).
71. Nikolova, T., Roos, W. P., Krämer, O. H., Strik, H. M. & Kaina, B. Chloroethylating nitrosoureas in cancer therapy: DNA damage, repair and cell death signaling. *Biochim. Biophys. Acta BBA - Rev. Cancer* **1868**, 29–39 (2017).
72. Pfeifer, G. P. *et al.* Tobacco smoke carcinogens, DNA damage and p53 mutations in smoking-associated cancers. *Oncogene* **21**, 7435–7451 (2002).
73. Boström, C.-E. *et al.* Cancer risk assessment, indicators, and guidelines for polycyclic aromatic hydrocarbons in the ambient air. *Environ. Health Perspect.* **110**, 451–488 (2002).
74. Hsiao, Y.-C., Liu, C.-W., Chi, L., Yang, Y. & Lu, K. Effects of Gut Microbiome on Carcinogenic DNA Damage. *Chem. Res. Toxicol.* **33**, 2130–2138 (2020).
75. Chung, F.-L., Chen, H.-J. C. & Nath, R. G. Lipid peroxidation as a potential endogenous source for the formation of exocyclic DNA adducts. *Carcinogenesis* **17**, 2105–2111 (1996).
76. Pullman, A. & Pullman, B. Molecular electrostatic potential of the nucleic acids. *Q. Rev. Biophys.* **14**, 289–380 (1981).
77. Beranek, D. T. Distribution of methyl and ethyl adducts following alkylation with monofunctional alkylating agents. *Mutat. Res. Mol. Mech. Mutagen.* **231**, 11–30 (1990).

78. Betteridge, D. J. What is oxidative stress? *Metab. - Clin. Exp.* **49**, 3–8 (2000).
79. Zuo, L., Zhou, T., Pannell, B. K., Ziegler, A. C. & Best, T. M. Biological and physiological role of reactive oxygen species – the good, the bad and the ugly. *Acta Physiol.* **214**, 329–348 (2015).
80. Brieger, K., Schiavone, S., Jr, F. J. M. & Krause, K.-H. Reactive oxygen species: from health to disease. *Swiss Med. Wkly.* **142**, 13659–13659 (2012).
81. van der Vliet, A., Janssen-Heininger, Y. M. W. & Anathy, V. Oxidative stress in chronic lung disease: From mitochondrial dysfunction to dysregulated redox signaling. *Mol. Aspects Med.* **63**, 59–69 (2018).
82. Huang, M.-F., Lin, W.-L. & Ma, Y.-C. A study of reactive oxygen species in mainstream of cigarette. *Indoor Air* **15**, 135–140 (2005).
83. Hung, H.-F. & Wang, C.-S. Experimental determination of reactive oxygen species in Taipei aerosols. *J. Aerosol Sci.* **32**, 1201–1211 (2001).
84. Heck, D. E., Vetrano, A. M., Mariano, T. M. & Laskin, J. D. UVB Light Stimulates Production of Reactive Oxygen Species. *J. Biol. Chem.* **278**, 22432–22436 (2003).
85. Azzam, E. I., Jay-Gerin, J.-P. & Pain, D. Ionizing radiation-induced metabolic oxidative stress and prolonged cell injury. *Cancer Lett.* **327**, 48–60 (2012).
86. Murphy, M. P. How mitochondria produce reactive oxygen species. *Biochem. J.* **417**, 1–13 (2008).
87. Winterbourn, C. C. Toxicity of iron and hydrogen peroxide: the Fenton reaction. *Toxicol. Lett.* **82–83**, 969–974 (1995).
88. Strlič, M., Kolar, J., Šelih, V., Kočar, D. & Pihlar, B. A comparative study of several transition metals in Fenton-like reaction systems at circum-neutral pH. *Acta Chim. Slov.* (2003).
89. Hofer, T. & Perry, G. Nucleic acid oxidative damage in Alzheimer’s disease—explained by the hepcidin-ferroportin neuronal iron overload hypothesis? *J. Trace Elem. Med. Biol.* **38**, 1–9 (2016).
90. Kehrer, J. P. The Haber–Weiss reaction and mechanisms of toxicity. *Toxicology* **149**, 43–50 (2000).

91. Yermilov, V., Rubio, J. & Ohshima, H. Formation of 8-nitroguanine in DNA treated with peroxyxynitrite *in vitro* and its rapid removal from DNA by depurination. *FEBS Lett.* **376**, 207–210 (1995).
92. Alshykhly, O. R., Fleming, A. M. & Burrows, C. J. 5-Carboxamido-5-formamido-2-iminohydantoin, in Addition to 8-oxo-7,8-Dihydroguanine, Is the Major Product of the Iron-Fenton or X-ray Radiation-Induced Oxidation of Guanine under Aerobic Reducing Conditions in Nucleoside and DNA Contexts. *J. Org. Chem.* **80**, 6996–7007 (2015).
93. O'Donnell, V. B. *et al.* Nitric Oxide Inhibition of Lipid Peroxidation: Kinetics of Reaction with Lipid Peroxyl Radicals and Comparison with α -Tocopherol. *Biochemistry* **36**, 15216–15223 (1997).
94. Aruoma, O. I., Halliwell, B. & Dizdaroglu, M. Iron Ion-dependent Modification of Bases in DNA by the Superoxide Radical-generating System Hypoxanthine/Xanthine Oxidase. *J. Biol. Chem.* **264**, 13024–13028 (1989).
95. Steenken, S., Jovanovic, S. V., Bietti, M. & Bernhard, K. The Trap Depth (in DNA) of 8-Oxo-7,8-dihydro-2'-deoxyguanosine as Derived from Electron-Transfer Equilibria in Aqueous Solution. *J. Am. Chem. Soc.* **122**, 2373–2374 (2000).
96. Matter, B. *et al.* Mapping three guanine oxidation products along DNA following exposure to three types of reactive oxygen species. *Free Radic. Biol. Med.* **121**, 180–189 (2018).
97. Vieira, A. J. S. C. & Steenken, S. Pattern of hydroxy radical reaction with adenine and its nucleosides and nucleotides. Characterization of two types of isomeric hydroxy adduct and their unimolecular transformation reactions. *J. Am. Chem. Soc.* **112**, 6986–6994 (1990).
98. Samson-Thibault, F., Madugundu, G. S., Gao, S., Cadet, J. & Wagner, J. R. Profiling Cytosine Oxidation in DNA by LC-MS/MS. *Chem. Res. Toxicol.* **25**, 1902–1911 (2012).
99. Wagner, J. R., Hu, C. C. & Ames, B. N. Endogenous oxidative damage of deoxycytidine in DNA. *Proc. Natl. Acad. Sci.* **89**, 3380–3384 (1992).
100. Rivière, J., Ravanat, J.-L. & Wagner, J. R. Ascorbate and H₂O₂ induced oxidative DNA damage in Jurkat cells. *Free Radic. Biol. Med.* **40**, 2071–2079 (2006).

101. Lenton, K. J., Therriault, H., Fülöp, T., Payette, H. & Wagner, J. R. Glutathione and ascorbate are negatively correlated with oxidative DNA damage in human lymphocytes. *Carcinogenesis* **20**, 607–613 (1999).
102. Chattopadhyaya, R. Oxidative damage to DNA constituents by iron-mediated Fenton reactions – the thymidine family. *J. Biomol. Struct. Dyn.* **32**, 155–169 (2014).
103. Rausch, C. *et al.* Cytosine base modifications regulate DNA duplex stability and metabolism. *Nucleic Acids Res.* **49**, 12870–12894 (2021).
104. Bielskutè, S., Plavec, J. & Podbevšek, P. Impact of Oxidative Lesions on the Human Telomeric G-Quadruplex. *J. Am. Chem. Soc.* **141**, 2594–2603 (2019).
105. Banda, D. M., Nuñez, N. N., Burnside, M. A., Bradshaw, K. M. & David, S. S. Repair of 8-oxoG:A mismatches by the MUTYH glycosylase: Mechanism, metals and medicine. *Free Radic. Biol. Med.* **107**, 202–215 (2017).
106. Fujii, S., Wang, A. H.-J., van der Marel, G., van Boom, J. H. & Rich, A. Molecular structure of (m⁵ dC-dG)₃ : the role of the methyl group on 5-methyl cytosine in stabilizing Z-DNA. *Nucleic Acids Res.* **10**, 7879–7892 (1982).
107. Mayer-Jung, C., Moras, D. & Timsit, Y. Hydration and recognition of methylated CpG steps in DNA. *EMBO J.* **17**, 2709–2718 (1998).
108. Raindlová, V. *et al.* Influence of major-groove chemical modifications of DNA on transcription by bacterial RNA polymerases. *Nucleic Acids Res.* **44**, 3000–3012 (2016).
109. Wang, A. H., Hakoshima, T., van der Marel, G., van Boom, J. H. & Rich, A. AT base pairs are less stable than GC base pairs in Z-DNA: the crystal structure of d(m⁵CGTAm⁵CG). *Cell* **37**, 321–331 (1984).
110. Zhou, G. & Ho, P. S. Stabilization of Z-DNA by demethylation of thymine bases: 1.3-ANG. single-crystal structure of d(m⁵CGUAm⁵CG). *Biochemistry* **29**, 7229–7236 (1990).
111. Mooers, B. H. M., Schroth, G. P., Baxter, W. W. & Ho, S. P. Alternating and Non-alternating dG-dC Hexanucleotides Crystallize as Canonical A-DNA. *J. Mol. Biol.* **249**, 772–784 (1995).

112. Tippin, D. B. & Sundaralingam, M. Nine polymorphic crystal structures of d(CCGGGCCCCGG), d(CCGGGCCm5CGG), d(Cm5CGGGCCm5CGG) and d(CCGGGCC(Br)5CGG) in three different conformations: effects of spermine binding and methylation on the bending and condensation of A-DNA. *J. Mol. Biol.* **267**, 1171–1185 (1997).
113. Heinemann, U. & Hahn, M. C-C-A-G-G-C-m5C-T-G-G. Helical fine structure, hydration, and comparison with C-C-A-G-G-C-C-T-G-G. *J. Biol. Chem.* **267**, 7332–7341 (1992).
114. Renciuik, D., Blacque, O., Vorlickova, M. & Spingler, B. Crystal structures of B-DNA dodecamer containing the epigenetic modifications 5-hydroxymethylcytosine or 5-methylcytosine. *Nucleic Acids Res.* **41**, 9891–9900 (2013).
115. Vargason, J. M., Henderson, K. & Ho, P. S. A crystallographic map of the transition from B-DNA to A-DNA. *Proc. Natl. Acad. Sci.* **98**, 7265–7270 (2001).
116. Vargason, J. M. & Ho, P. S. The Effect of Cytosine Methylation on the Structure and Geometry of the Holliday Junction. *J. Biol. Chem.* **277**, 21041–21049 (2002).
117. Shi, K., Pan, B., Tippin, D. & Sundaralingam, M. Structures of d(Gm5CGm5CGCGC) and d(GCGCGm5CGm5C): effects of methylation on alternating DNA octamers. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 61–65 (2004).
118. Hays, F. A., Jones, Z. J. R. & Ho, P. S. Influence of Minor Groove Substituents on the Structure of DNA Holliday Junctions. *Biochemistry* **43**, 9813–9822 (2004).
119. Lercher, L. *et al.* Structural insights into how 5-hydroxymethylation influences transcription factor binding. *Chem. Commun.* **50**, 1794–1796 (2014).
120. Theruvathu, J. A., Yin, Y. W., Pettitt, B. M. & Sowers, L. C. Comparison of the Structural and Dynamic Effects of 5-Methylcytosine and 5-Chlorocytosine in a CpG Dinucleotide Sequence. *Biochemistry* **52**, 8590–8598 (2013).
121. Vander Zanden, C. M., Rowe, R. K., Broad, A. J., Robertson, A. B. & Ho, P. S. Effect of Hydroxymethylcytosine on the Structure and Stability of Holliday Junctions. *Biochemistry* **55**, 5781–5789 (2016).
122. Gruber, D. R. *et al.* Oxidative damage to epigenetically methylated sites affects DNA stability, dynamics and enzymatic demethylation. *Nucleic Acids Res.* **46**, 10827–10839 (2018).

123. Wan, L., Yi, J., Lam, S. L., Lee, H. K. & Guo, P. 5-Methylcytosine Substantially Enhances the Thermal Stability of DNA Minidumbbells. *Chem. – Eur. J.* **27**, 6740–6747 (2021).
124. Basti, M. M., Stuart, J. W., Lam, A. T., Guenther, R. & Agris, P. F. Design, biological activity and NMR-solution structure of a DNA analogue of yeast tRNA^{Phe} anticodon domain. *Nat. Struct. Biol.* **3**, 38–44 (1996).
125. Ramakrishnan, B. & Sundaralingam, M. Crystal structure of the A-DNA decamer d(CCIGGCCm5CGG) at 1.6 Å showing the unexpected wobble I.m5C base pair. *Biophys. J.* **69**, 553–558 (1995).
126. Shatzky-Schwartz, M. *et al.* X-ray and solution studies of DNA oligomers and implications for the structural basis of A-tract-dependent curvature. *J. Mol. Biol.* **267**, 595–623 (1997).
127. Mayer-Jung, C., Moras, D. & Timsit, Y. Effect of cytosine methylation on DNA-DNA recognition at CpG steps¹¹Edited by T. Richmond. *J. Mol. Biol.* **270**, 328–335 (1997).
128. Tippin, D. B., Ramakrishnan, B. & Sundaralingam, M. Methylation of the Z-DNA decamer d(GC)₅ potentiates the formation of A-DNA: crystal structure of d(Gm5CGm5CGCGCGC)₁₁Edited by I. Tinoco. *J. Mol. Biol.* **270**, 247–258 (1997).
129. Mooers, B. H. M., Eichman, B. F. & Ho, P. S. The structures and relative stabilities of d(G · G) reverse Hoogsteen, d(G · T) reverse wobble, and d(G · C) reverse Watson-Crick base-pairs in DNA crystals¹¹Edited by I. Tinoco. *J. Mol. Biol.* **269**, 796–810 (1997).
130. Eichman, B. F., Schroth, G. P., Basham, B. E. & Ho, P. S. The intrinsic structure and stability of out-of-alternation base pairs in Z-DNA. *Nucleic Acids Res.* **27**, 543–550 (1999).
131. Phan, A. T., Guéron, M. & Leroy, J.-L. The solution structure and internal motions of a fragment of the cytidine-rich strand of the human telomere. *J. Mol. Biol.* **299**, 123–144 (2000).
132. Vargason, J. M., Eichman, B. F. & Ho, P. S. The extended and eccentric E-DNA structure induced by cytosine methylation or bromination. *Nat. Struct. Biol.* **7**, 758–761 (2000).
133. Mauffret, O. *et al.* A two B–Z junction containing DNA resolves into an all right-handed double-helix. *Nucleic Acids Res.* **28**, 4403–4409 (2000).

134. Wanunu, M. *et al.* Discrimination of Methylcytosine from Hydroxymethylcytosine in DNA Molecules. *J. Am. Chem. Soc.* **133**, 486–492 (2011).
135. Szulik, M. W. *et al.* Differential Stabilities and Sequence-Dependent Base Pair Opening Dynamics of Watson–Crick Base Pairs with 5-Hydroxymethylcytosine, 5-Formylcytosine, or 5-Carboxylcytosine. *Biochemistry* **54**, 1294–1305 (2015).
136. Raiber, E.-A. *et al.* 5-Formylcytosine alters the structure of the DNA double helix. *Nat. Struct. Mol. Biol.* **22**, 44–49 (2015).
137. Wang, R. *et al.* Base pairing and structural insights into the 5-formylcytosine in RNA duplex. *Nucleic Acids Res.* **44**, 4968–4977 (2016).
138. Hardwick, J. S. *et al.* 5-Formylcytosine does not change the global structure of DNA. *Nat. Struct. Mol. Biol.* **24**, 544–552 (2017).
139. Frederick, C. A. *et al.* Methylation of the EcoRI recognition site does not alter DNA conformation: the crystal structure of d(CGCGAm6ATTCGCG) at 2.0-Å resolution. *J. Biol. Chem.* **263**, 17872–17879 (1988).
140. Baikalov, I., Grzeskowiak, K., Yanagi, K., Quintana, J. & Dickerson, R. E. The Crystal Structure of the Trigonal Decamer C-G-A-T-C-G-6meA-T-C-G: a B-DNA Helix with 10·6 Base-pairs Per Turn. *J. Mol. Biol.* **231**, 768–784 (1993).
141. Bae, S.-H. *et al.* Structure and Dynamics of Hemimethylated GATC Sites. *J. Biol. Chem.* **278**, 45987–45993 (2003).
142. Bang, J., Bae, S.-H., Park, C.-J., Lee, J.-H. & Choi, B.-S. Structural and Dynamics Study of DNA Dodecamer Duplexes That Contain Un-, Hemi-, or Fully Methylated GATC Sites. *J. Am. Chem. Soc.* **130**, 17688–17696 (2008).
143. Dunn, D. B. & Smith, J. D. Occurrence of a New Base in the Deoxyribonucleic Acid of a Strain of Bacterium Coli. *Nature* **175**, 336–337 (1955).
144. Holbrook, S. R., Dickerson, R. E. & Kim, S.-H. Anisotropic thermal-parameter refinement of the DNA dodecamer CGCGAATTCGCG by the segmented rigid-body method. *Acta Crystallogr. B* **41**, 255–262 (1985).

145. Michaels, M. L., Pham, L., Cruz, C. & Miller, J. H. MutM, a protein that prevents G C→T A transversions, is formamidopyrimidine-DNA glycosylase. *Nucleic Acids Res.* **19**, 3629–3632 (1991).
146. Lipscomb, L. A. *et al.* X-ray structure of a DNA decamer containing 7,8-dihydro-8-oxoguanine. *Proc. Natl. Acad. Sci.* **92**, 719–723 (1995).
147. McAuley-Hecht, K. E. *et al.* Crystal Structure of a DNA Duplex Containing 8-Hydroxydeoxyguanine-Adenine Base Pairs. *Biochemistry* **33**, 10266–10270 (1994).
148. Thiviyathan, V., Somasunderam, A., Hazra, T. K., Mitra, S. & Gorenstein, D. G. Solution Structure of a DNA Duplex Containing 8-Hydroxy-2'-Deoxyguanosine Opposite Deoxyguanosine. *J. Mol. Biol.* **325**, 433–442 (2003).
149. Zálešák, J., Constant, J.-F. & Jourdan, M. Nuclear Magnetic Resonance Solution Structure of DNA Featuring Clustered 2'-Deoxyribonolactone and 8-Oxoguanine Lesions. *Biochemistry* **55**, 3899–3906 (2016).
150. Zálešák, J., Lourdin, M., Krejčí, L., Constant, J.-F. & Jourdan, M. Structure and Dynamics of DNA Duplexes Containing a Cluster of Mutagenic 8-Oxoguanine and Abasic Site Lesions. *J. Mol. Biol.* **426**, 1524–1538 (2014).
151. Hoppins, J. J. *et al.* 8-Oxoguanine Affects DNA Backbone Conformation in the EcoRI Recognition Site and Inhibits Its Cleavage by the Enzyme. *PLOS ONE* **11**, e0164424 (2016).
152. Saito, I. *et al.* Mapping of the Hot Spots for DNA Damage by One-Electron Oxidation: Efficacy of GG Doublets and GGG Triplets as a Trap in Long-Range Hole Migration. *J. Am. Chem. Soc.* **120**, 12686–12687 (1998).
153. Bielskutè, S., Plavec, J. & Podbevšek, P. Oxidative lesions modulate G-quadruplex stability and structure in the human BCL2 promoter. *Nucleic Acids Res.* **49**, 2346–2356 (2021).
154. Cheong, V. V., Heddi, B., Lech, C. J. & Phan, A. T. Xanthine and 8-oxoguanine in G-quadruplexes: formation of a G·G·X·O tetrad. *Nucleic Acids Res.* **43**, 10506–10514 (2015).
155. Leonard, G. A., Guy, A., Brown, T., Teoule, R. & Hunter, W. N. Conformation of Guanine-8-Oxoadenine Base Pairs in the Crystal Structure of d(CGCGAATT(08A)GCG). *Biochemistry* **31**, 8415–8420 (1992).

156. Chen, H., Johnson, F., Grollman, A. P. & Patel, D. J. NMR Study of an 11-Mer DNA Duplex Containing 7,8-Dihydro-8-Oxoadenine (AOXO) Opposite Thymine. <https://www.rcsb.org/structure/1fjb> (2015).
157. Sugiyama, H. *et al.* Synthesis, Structure and Thermodynamic Properties of 8-Methylguanine-Containing Oligonucleotides: Z-DNA under Physiological Salt Conditions. *Nucleic Acids Res.* **24**, 1272–1278 (1996).
158. Zhang, F. *et al.* Structures of DNA duplexes containing O6-carboxymethylguanine, a lesion associated with gastrointestinal cancer, reveal a mechanism for inducing pyrimidine transition mutations. *Nucleic Acids Res.* **41**, 5524–5532 (2013).
159. Zhang, F. *et al.* O6-Carboxymethylguanine in DNA forms a sequence context-dependent wobble base-pair structure with thymine. *Acta Crystallogr. D Biol. Crystallogr.* **70**, 1669–1679 (2014).
160. Lloyd, D. R. & Phillips, D. H. Oxidative DNA damage mediated by copper(II), iron(II) and nickel(II) Fenton reactions: evidence for site-specific mechanisms in the formation of double-strand breaks, 8-hydroxydeoxyguanosine and putative intrastrand cross-links. *Mutat. Res. Mol. Mech. Mutagen.* **424**, 23–36 (1999).
161. Imlay, J. A., Chin, S. M. & Linn, S. Toxic DNA Damage by Hydrogen Peroxide Through the Fenton Reaction *in Vivo* and *in Vitro*. *Science* **240**, 640–642 (1988).
162. Halliwell, B., Adhikary, A., Dingfelder, M. & Dizdaroglu, M. Hydroxyl radical is a significant player in oxidative DNA damage *in vivo*. *Chem. Soc. Rev.* **50**, 8355–8360 (2021).
163. Kazakov, S. A. & Hecht, S. M. Nucleic Acid–Metal Ion Interactions. in *Encyclopedia of Inorganic Chemistry* (John Wiley & Sons, Ltd, 2006).
164. Abrescia, N. G. A. *et al.* Structure of the oligonucleotide d(CGTATATACG) as a site-specific complex with nickel ions. *Nucleic Acids Res.* **27**, 1593–1599 (1999).
165. Abrescia, N. G. A., Malinina, L. & Subirana, J. A. Stacking interaction of guanine with netropsin in the minor groove of d(CGTATATACG)₂. *J. Mol. Biol.* **294**, 657–666 (1999).
166. Abrescia, N. G., Huynh-Dinh, T. & Subirana, J. A. Nickel-guanine interactions in DNA: crystal structure of nickel-d[CGTGTACACG]₂. *JBIC J. Biol. Inorg. Chem.* **7**, 195–199 (2002).

167. Labiuk, S. L., Delbaere, L. T. J. & Lee, J. S. Cobalt(II), nickel(II) and zinc(II) do not bind to intra-helical N(7) guanine positions in the B-form crystal structure of d(GGCGCC). *JBIC J. Biol. Inorg. Chem.* **8**, 715–720 (2003).
168. Valls, N., Usón, I., Gouyette, C. & Subirana, J. A. A Cubic Arrangement of DNA Double Helices Based on Nickel–Guanine Interactions. *J. Am. Chem. Soc.* **126**, 7812–7816 (2004).
169. Tseng, W.-H. *et al.* Induced-Fit Recognition of CCG Trinucleotide Repeats by a Nickel–Chromomycin Complex Resulting in Large-Scale DNA Deformation. *Angew. Chem. Int. Ed.* **56**, 8761–8765 (2017).
170. Jhan, C.-R. *et al.* Targeting the ALS/FTD-associated A-DNA kink with anthracene-based metal complex causes DNA backbone straightening and groove contraction. *Nucleic Acids Res.* **49**, 9526–9538 (2021).
171. Valls, N., Wright, G., Steiner, R. A., Murshudov, G. N. & Subirana, J. A. DNA variability in five crystal structures of d(CGCAATTGCG). *Acta Crystallogr. D Biol. Crystallogr.* **60**, 680–685 (2004).
172. Adams, A., Guss, J. M., Denny, W. A. & Wakelin, L. P. G. Structure of 9-amino-[N-(2-dimethylamino)propyl]acridine-4-carboxamide bound to d(CGTACG)₂: a comparison of structures of d(CGTACG)₂ complexed with intercalators in the presence of cobalt. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 823–828 (2004).
173. Valls, N., Steiner, R. A., Wright, G., Murshudov, G. N. & Subirana, J. A. Variable role of ions in two drug intercalation complexes of DNA. *JBIC J. Biol. Inorg. Chem.* **10**, 476–482 (2005).
174. Drozdal, P., Gilski, M., Kierzek, R., Lomozik, L. & Jaskolski, M. Ultrahigh-resolution crystal structures of Z-DNA in complex with Mn²⁺ and Zn²⁺ ions. *Acta Crystallogr. D Biol. Crystallogr.* **69**, 1180–1190 (2013).
175. Hendrickson, W. A. Determination of Macromolecular Structures from Anomalous Diffraction of Synchrotron Radiation. *Science* **254**, 51–58 (1991).
176. Hendrickson, W. A. & Teeter, M. M. Structure of the hydrophobic protein crambin determined directly from the anomalous scattering of sulphur. *Nature* **290**, 107–113 (1981).

177. Ennifar, E., Walter, P. & Dumas, P. An efficient method for solving RNA structures: MAD phasing by replacing magnesium with zinc. *Acta Crystallogr. D Biol. Crystallogr.* **57**, 330–332 (2001).
178. Hou, C. & Tsodikov, O. V. Utilizing guanine-coordinated Zn²⁺ ions to determine DNA crystal structures by single-wavelength anomalous diffraction. *Acta Crystallogr. Sect. D* **75**, 32–40 (2019).
179. Berry, K. E., Waghray, S., Mortimer, S. A., Bai, Y. & Doudna, J. A. Crystal Structure of the HCV IRES Central Domain Reveals Strategy for Start-Codon Positioning. *Structure* **19**, 1456–1466 (2011).
180. Adams, A., Guss, J. M., Collyer, C. A., Denny, W. A. & Wakelin, L. P. G. A novel form of intercalation involving four DNA duplexes in an acridine-4-carboxamide complex of d(CGTACG)₂. *Nucleic Acids Res.* **28**, 4244–4253 (2000).
181. Chen, Y.-W., Jhan, C.-R., Neidle, S. & Hou, M.-H. Structural Basis for the Identification of an i-Motif Tetraplex Core with a Parallel-Duplex Junction as a Structural Motif in CCG Triplet Repeats. *Angew. Chem. Int. Ed.* **53**, 10682–10686 (2014).
182. Simmons, C. R. *et al.* The influence of Holliday junction sequence and dynamics on DNA crystal self-assembly. *Nat. Commun.* **13**, 3112 (2022).
183. Matoba, Y., Kumagai, T., Yamamoto, A., Yoshitsu, H. & Sugiyama, M. Crystallographic Evidence That the Dinuclear Copper Center of Tyrosinase Is Flexible during Catalysis. *J. Biol. Chem.* **281**, 8981–8990 (2006).
184. Porta, J., Vahedi-Faridi, A. & Borgstahl, G. E. O. Structural Analysis of Peroxide-Soaked MnSOD Crystals Reveals Side-On Binding of Peroxide to Active-Site Manganese. *J. Mol. Biol.* **399**, 377–384 (2010).
185. Bailey, L. J. & Fox, B. G. Crystallographic and Catalytic Studies of the Peroxide-Shunt Reaction in a Diiron Hydroxylase. *Biochemistry* **48**, 8932–8939 (2009).
186. Marangon, J. *et al.* Kinetic and Structural Studies of Aldehyde Oxidoreductase from *Desulfovibrio gigas* Reveal a Dithiolene-Based Chemistry for Enzyme Activation and Inhibition by H₂O₂. *PLOS ONE* **8**, e83234 (2013).

187. Peskin, A. V. *et al.* Modifying the resolving cysteine affects the structure and hydrogen peroxide reactivity of peroxiredoxin 2. *J. Biol. Chem.* **296**, (2021).
188. Nehls, T., Heymann, T., Meyners, C., Hausch, F. & Lermyte, F. Fenton-Chemistry-Based Oxidative Modification of Proteins Reflects Their Conformation. *Int. J. Mol. Sci.* **22**, 9927 (2021).
189. Shalev, D. E. Studying Peptide-Metal Ion Complex Structures by Solution-State NMR. *Int. J. Mol. Sci.* **23**, 15957 (2022).

Chapter 2 - Co-crystallization vs Soaking - a Comparison of Methodologies for Obtaining Ordered Metal Sites in Crystals with Minimum Structural Perturbation

2.1. Introduction

Metal ions are necessary for many biological processes¹. 30 to 40 % of proteins need at least one metal ion to perform their functions², whilst for nucleic acids, metal ions were determined to assist structural stability³ and mediate nucleic acid-protein interactions⁴. Generally, alkali and alkaline-earth metals, mostly Na⁺, K⁺, Mg²⁺, or Ca²⁺, have been shown to influence both nucleic acids⁵⁻⁷ and proteins⁸ structures. Transition metals, such as Co²⁺, Cu²⁺, Fe²⁺, Zn²⁺ or Mn²⁺, however, were shown to both influence structures⁹ and have catalytic roles^{10,11}. Consequently, structural methods have been the preferred approaches to obtain critical information and details on macromolecules-metal interactions and gain insight on the role of metal ions in cellular environments. One of these methods, X-ray crystallography, accounts for ~85 % of all published atomic coordinates deposited in the Protein Data Base¹² (PDB). As such, metal binding sites have been frequently reported in biological macromolecular structures deposited in the database, and as of May 2022, out of the ~150,000 structures listed on metalPDB¹³, 38 % were metal-bound.

Transition metals have also been used in X-ray crystallography for their usefulness in solving the “phase problem” originating from the very nature of a diffraction experiment: how can the electron density map be calculated if only intensities are collected, and the phases are lost¹⁴? The electron density at a (x, y, z) position in the real space can be calculated by using the Fourier transform (Scheme 2.1.(1)) and by summing up all the contributions to the (x, y, z) position of a wave scattered from a plan (hkl). The structure factors F_{hkl} can be described by both its amplitude $|F_{hkl}|$, which is the square root of the intensity I_{hkl} of all electrons in the (hkl) plane, and its associated phase α_{hkl} (2) to give equation (3)¹⁵.

$$(1) \rho(x, y, z) = \frac{1}{V} \sum_{hkl} F_{hkl} e^{-2\pi(hx+ky+lz)}$$

$$(2) F_{hkl} = |F_{hkl}| e^{i\alpha_{hkl}}$$

$$(3) \rho(x, y, z) = \frac{1}{V} \sum_{hkl} |F_{hkl}| e^{i\alpha_{hkl}} e^{-2\pi(hx+ky+lz)}$$

Scheme 2.1: Electron density and structure factors equations calculated using the Fourier transform

Phases cannot be calculated on the basis of intensity information alone and the only way to retrieve the missing phases would be by having hold of the electron density map already, which is not available¹⁵. However, with ~200,000 structures in the PDB available, and with prior knowledge of the investigated structure, a starting model can be used to calculate an initial set of phases, provided that this model has a low Root Mean Square Deviation (RMSD) of atomic positions with the unsolved structure, a method called Molecular Replacement¹⁶. The majority of X-ray crystal structures are solved using this technique. However, if this method fails, other methods have been developed and make use of heavy atoms to find initial phases. The process of incorporating these heavy atoms to a native crystal is called derivatisation, and many novel structures are solved thereafter using either Single and Multiple Isomorphous Replacement (SIR/MIR)¹⁷ or Single and Multi-wavelength Anomalous Dispersion (SAD/MAD)^{18,19}. SIR and MIR require both a native crystal and a derivative crystal to obtain an estimate of the heavy atom structure factor and determine the heavy atoms' positions using either the direct method or the Patterson function to get approximate phases. SAD and MAD, however, do not require a native system and only the knowledge of the correct wavelength, at which an electron from the heavy atom's inner shell can be promoted. Under these anomalous scattering conditions, the structure factor of the heavy atom f can also be described by equation (4), with two correction terms, the real number f' and the imaginary component f'' , which are wavelength dependent. This results in Friedel's Law, which states that members of a Friedel pair F_{hkl} and $F_{\bar{h}\bar{k}\bar{l}}$ have equal amplitudes and opposite phase, to be broken (Scheme 2.2.(5)).

$$(4) f = f_0 + f'(\lambda) + if''(\lambda)$$

$$(5) |F_{hkl}| \neq |F_{\bar{h}\bar{k}\bar{l}}|$$

Scheme 2.2: Structure factor equation of a heavy atom and broken Friedel's Law

This is sufficient to obtain the heavy atom anomalous contributions, and by drawing a Harker construction using the Argand diagram, two possible solutions for the phase of each structure factor can be determined (Figure 2.1).

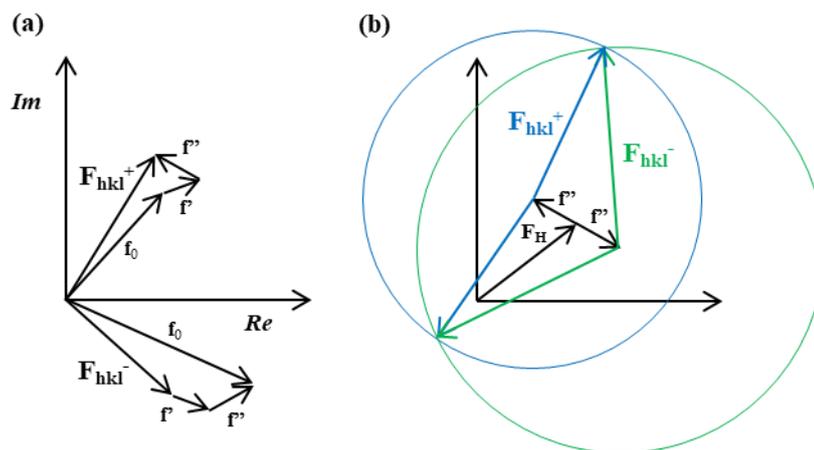


Figure 2.1: a. Separation of the f_0 , f' and f'' components of a Bijvoet pair of structure factors presented on an Argand diagram. b. The Harker construction for the SAD method.

The addition of transition-metal ions to crystal systems containing biomolecules can generally be achieved in two ways:

1. Crystallisation from a solution containing both the metal ion and the biomolecules, which can be referred to as cocrystallisation
2. Soaking a grown crystal with a solution containing the metal ion of interest

Both methods have advantages and disadvantages, so it is important that an appropriate method be chosen to match the requirements of the study and the proposed outcomes. The first method, if successful, will likely give the thermodynamically favoured structure, however the process of growing crystals is long, and some transition metals might not be stable at certain pH²⁰, oxidise easily in aerobic environment²¹ or precipitate the macromolecules²². The second method exploits the solvent channels present in both protein and nucleic acid crystals. On average, about 50 % of the unit-cell is filled with the mother liquor, an estimate originally

observed by Matthews²³. Yet, the distribution in solvent content is wide, ranging from 26 % to 90 %, and the size of solvent channel might not be correlated to solvent volume. Nonetheless, metal ions can diffuse through and hopefully give ordered binding sites. This is the more commonly applied procedure, especially in cases when cocrystallisation is not possible for the reasons listed above. Yet, it is important to estimate both accurate concentration for the solution, as a too high concentration can damage the crystal through osmotic shock²⁴, and soaking times to fully populate the crystal, as short soaking times might lead to the centre and superficial regions of the crystal to be populated differently²⁵. However, while both methods *could* be used, it is important to know if there are cases where the choice of methodology could influence the observed biological structure. If so, then it may be the case that one method is preferred over the other and can therefore be relied on to obtain a structure which is more representative of that in a biological environment.

Protein-ligand systems have been investigated to try and answer that question^{26,27}, however, the effect of the presence of transition metals on both methods on nucleic acid structures, and more particularly DNA structures, have not been investigated. This is of even greater importance since metal-mediated oxidative DNA damage has been associated with many diseases, and knowledge of where metal binding sites are could bring more insight on the location of metal reaction centres as well as educate us on which bases are more susceptible to damage thanks to structural information. In the long-term, this could help predict hotspots centres for damage with prior knowledge of the sequence and the structure this sequence can potentially form. Working with both iron and copper would be highly preferable, however, to prevent iron from undergoing oxidation, anoxic conditions are needed, while DNA precipitates in the presence of copper at the concentrations employed for cocrystallisation. Thus, there is a need to use other transition metals, such as nickel and cobalt, as model systems.

The DNA sequence d(GCATGCT) has previously been reported by the Cardin group to crystallise in a two-stranded non-classical quadruplex form²⁸ in the presence of vanadium, barium, and calcium (Figure 2.2) and in a four-stranded non-classical quadruplex in the presence of cobalt and nickel.

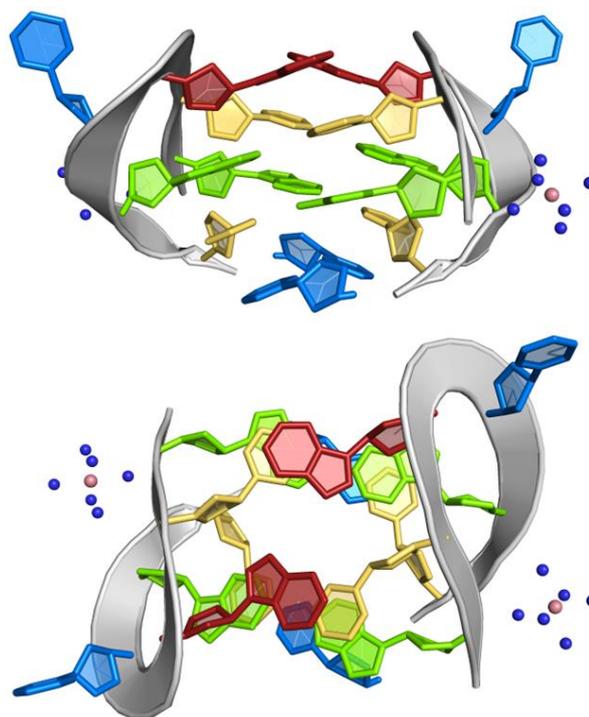


Figure 2.2: Biological unit of the 1MF5²⁵ structure, a non-classical quadruplex, with views from the side (top) and from the top (bottom). Nucleotides are coloured following convention, guanine bases in green, adenine bases in red and thymine bases in blue. The backbone chain is in grey. Hexamine cobalt (III) ions are represented as spheres.

Here, two structures of d(GCATGCT) containing bound Ni²⁺ and Co²⁺ ions, obtained using the soaking method, are reported. The results, the native structure and the structures of the same sequence grown using cocrystallisation reported in the PDB are compared to each other to identify which method causes the less perturbation of the structure, especially when dealing with non-canonical structures, as this could prove meaningful in the comprehension of metal-mediated oxidative DNA damage and to develop new techniques to study these types of reactions in a crystalline environment.

2.2. Material and methods

2.2.1. DNA preparation

The oligonucleotide d(GCATGCT) was purchased from Eurogentec as a solid purified by RP-HPLC. DNA concentration was determined from the value of the absorbance at $\lambda = 260$ nm, using the molar extinction coefficient provided by the manufacturer. 0.5 mM solution of the quadruplex forming sequence was annealed in a 40 mM sodium cacodylate trihydrate pH 5.7 buffer at 90°C for two minutes then allowed to cool slowly to room temperature.

2.2.2. Crystallisation and metal-soaking

Single crystals to be used for soaking with metal ions were grown at 291 K by the sitting-drop vapor diffusion method using the protocol described by Thorpe *et al*²⁸. The drops contained 2 μ L of 0.5 mM oligonucleotides and 2 μ L of a crystallisation condition containing 2 % (v/v) 2-methyl-2,4-pentanediol (MPD), 9 mM sodium cacodylate trihydrate pH 5.5, 18 mM hexamine cobalt (III), 7 mM sodium chloride and 178 mM potassium chloride. The drops were equilibrated against 500 μ l of 35 % (v/v) MPD. Orange crystals grew within four months (Figure 2.3) and were used for soaking in a drop containing 2 μ L of the crystallisation conditions and 2 μ L of either 10 mM NiCl₂ or 10 mM CoCl₂, for 4 hours and 3 days respectively. During nickel-soaking, the crystals' orange colour appeared to fade into an almost translucent blue colour (Figure 2.3). This change was used to track the experiment.

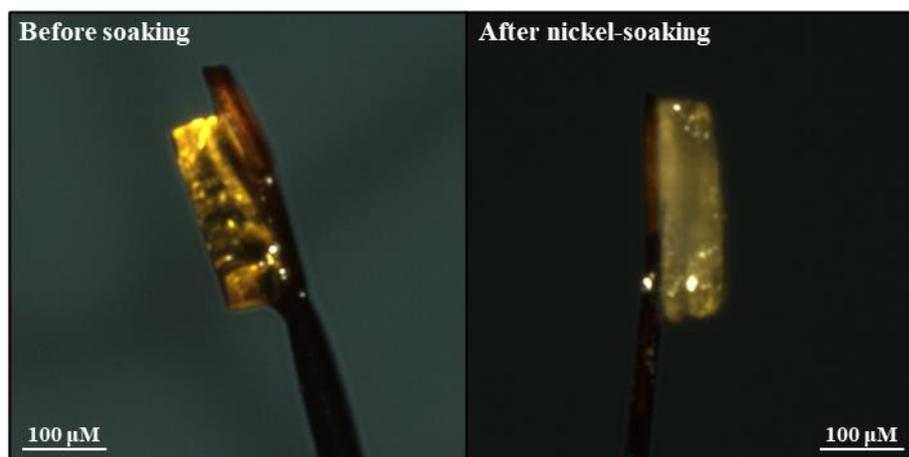


Figure 2.3: Crystals of the DNA sequence d(GCATGCT) before (left) and after (right) the crystals were soaked in a drop containing 10 mM NiCl₂. During the soaking step, the crystals' orange colour would fade into an almost translucent blue colour. This change was used to monitor the experiment.

2.2.3. Data collection and processing

Single crystals of approximate dimensions 200 x 100 x 20 μm were flash cooled with liquid nitrogen after soaking.

A data set from a native crystal prior to soaking was collected to the resolution of 1.08 Å at 100 K with 0.004s frames on beam line I03 at Diamond Light Source Ltd using radiation with a wavelength of 0.8266Å. 3600 frames were collected with an oscillation angle of 0.1° to give 360° of data. The data were integrated and scaled using xia2²⁹ with dials³⁰ giving 9890 unique reflections. The structure was first solved in *C*222₁ by molecular replacement using Phaser³¹. A model of the expected structure, **1MF5** from the Protein Data Bank, was selected to use for molecular replacement, used to provide initial estimate of the phases for this data. During molecular replacement, all possible orientations and positions of the model are tested against the data and calculate an initial set of phases. However, to confirm that the model is appropriate and will fit the data, the Matthews coefficient was calculated as it can indicate how many molecules are in the asymmetric unit based on the solvent content of the crystal^{23,32}. In this case, two strands of the **1MF5** model were used. The model was then updated using Coot³³ and refined using phenix.refine³⁴. To confirm that the model was satisfactory and that the observed data and the calculated data from the model were in agreement, the R_{free} and the R_{work} were calculated after each round of refinement. R-factors can be made arbitrarily low by overfitting the data during refinement. Consequently, five percent of the reflections were

reserved for the R_{free} set and not used during refinement³⁵. While a R_{free} of 0 % would mean total agreement between the model and the data and a R_{free} of upwards of 40 % would suggest the model is not a solution. In practice, the R_{free} is positively correlated with resolution and should be put in context with other structures of similar resolution³⁶. The R_{free} and R_{work} observed after the final round of refinement for this model were 0.1850 of 0.1638 respectively. The data, and final coordinates, were deposited in the Protein Data Bank with ID **8OE3**. The data collection and refinement statistics can be found in Table 2.1.

A data set from the nickel-soaked crystal was collected to the resolution of 1.19 Å at 100 K with 0.004s frames on beam line I03 at Diamond Light Source Ltd using radiation with a wavelength of 0.9763Å. 3600 frames were collected with an oscillation angle of 0.1° to give 360° of data. The data were integrated and scaled using xia2²⁹ with dials³⁰ giving 9890 unique reflections. The structure was first solved in P1 by molecular replacement using Phaser²⁸ and the structure 1MF5 from the Protein Data Bank, as MR with the 1MF5 model was first unsuccessful in $C222_1$. Eight strands were successfully placed in the P1 model. Two of these eight strands were then used as a suitable model to solve the structure in $C222_1$ using MR. The model was updated using Coot³³ and refined using phenix.refine³⁴ to give a final R_{free} of 0.1771 and R_{work} of 0.1447. Five percent of reflections were reserved for the R_{free} set. The data, and final coordinates, were deposited in the Protein Data Bank with ID **8ASO**. The data collection and refinement statistics for 8ASO can be found in Table 2.1.

Data for the cobalt-soaked crystal were collected operating a XtaLAB Synergy-S diffractometer using a PhotonJet-S X-ray source at a wavelength of 1.5406 Å with a HyPix-6000HE detector. 676 frames were collected. The data were integrated and scaled using the CrysAlisPRO software, which uses the CCP4 suite^{37,37}, giving 3466 unique reflections. The structure was also solved in $C222_1$ by molecular replacement using Phaser³¹ and the structure 1MF5 from the Protein Data Bank. The model was updated using Coot³³ and refined using phenix.refine³⁴ to give a final R_{free} of 0.2000 and R_{work} of 0.1784. Five percent of reflections were reserved for the R_{free} set. The data, and final coordinates, were deposited in the Protein Data Bank with ID **8ASM**. The data collection and refinement statistics for 8ASM can be found in Table 2.1.

Table 2.1: Data collection and refinement statistics for 8OE3, 8ASO and 8ASM. Outer shell statistics are shown in parentheses.

PDB ID	8OE3	8ASO	8ASM
Data collection			
Radiation source	I03 - DLS	I03 - DLS	XtaLAB Synergy-S diffractometer
Wavelength (Å)	0.8266	0.9763	1.5406
Space group	C222 ₁	C222 ₁	C222 ₁
Cell dimensions <i>a</i> , <i>b</i> , <i>c</i> , (Å)	22.16, 59.16, 45.67	21.95, 58.97, 45.60	21.90, 58.66, 45.57
Resolution, (Å)	24.83 - 1.08 (1.10 - 1.08)	22.80 - 1.19 (1.21 - 1.19)	18.71 - 1.70 (1.75 - 1.70)
R _{merge}	0.074 (0.396)	0.054 (0.148)	0.042 (0.103)
R _{meas}	0.081 (0.453)	0.056 (0.155)	0.047 (0.124)
R _{pim}	0.032 (0.214)	0.016 (0.043)	0.021 (0.059)
Total number of observations	82689	125551	31743
Total number of unique observations	13221	9890	3466
CC _{1/2}	0.998 (0.893)	1.000 (0.997)	0.999 (0.994)
I/σI	14.0 (2.0)	26.8 (7.5)	30.1 (13.5)
Completeness, %	99.6 (91.8)	100.0 (99.8)	99.9 (100.0)
Multiplicity	6.3 (4.1)	12.7 (12.7)	9.2 (5.6)
Refinement			
Resolution, (Å)	24.825 - 1.080 (1.119 - 1.080)	22.799 - 1.190 (1.232 - 1.190)	18.707 - 1.700 (1.761 - 1.700)
No. Reflections	13176	9853	3438
R _{work} /R _{free}	0.1638/0.1850	0.1447/0.1771	0.1784/0.2000
rmsd			
Bond lengths, (Å)	0.013	0.012	0.009
Bond angles, (°)	1.71	1.34	0.98

2.2.4. Structural and packing analysis

All figures were created using PyMOL³⁸ unless specified otherwise. Dinucleotide step parameters were calculated using Web 3DNA 2.0³⁹ are available in the Supplementary Information. CCP4 maps were generated from the final mtz file, including the anomalous map when available, using the phenix.mtz2map program. Maps were visualised in PyMOL and displayed using the mesh function. *2mFo-DFc* maps were contoured at 1σ unless specified otherwise. Anomalous maps were contoured at 3σ unless specified otherwise. The surface maps generated were visualised in PyMOL and displayed using the mesh function at level 2.0. Packing diagrams were generated using Mercury⁴⁰. Matthews' coefficients were calculated using the matthews_coef program found in the CCP4 suite³².

2.3. Results

2.3.1. Overall structure analysis - soaking experiment

8OE3 was obtained from a native crystal of the DNA heptamer sequence d(GCATGCT) (Figure 2.4). The crystal diffracted to a high resolution of 1.08 Å after collection on I03 at Diamond Light Source. This structure can be directly compared to **8ASM**, **8ASO** and **1MF5**²⁸.

The asymmetric unit of **8OE3** contains two strands for a total of 14 nucleotides, each of which forms a loop. Each strand forms a non-canonical quadruplex with a symmetry mate (Figure 2.4a). The nucleotides from strand A are labelled from G1 to T7 in the 5'-3' direction and the nucleotides from strand B are labelled from G8 to T14 in the 5'-3' direction. Nucleotides from symmetry related strands are designated with a * (Figure 2.4b). Disorder is only observed at the C13-T14 phosphate linkage. The alternative conformations are referred to as *a* and *b* and have occupancies of 60/40 % each. Details of the local base-pair parameters can be found in the Supplementary Information. The internal structure has been found to be identical to **1MF5**. Watson-Crick inter-strand bonds between oxoG1*-C6, C6*-oxoG1, C2*-G5 and G5*-C2 form the two G/C quadruplex formations, while Hoogsteen base pairs between A3-A3* and intra-strands bonds between G1-C6 and G5-C2 further stabilise the whole structure. The thymine bases form single-base loops which stacks with other thymine bases of

adjacent units. However, six hexamine cobalt residues were found in the asymmetric unit compared to the one hexamine cobalt residue in **1MF5** (Figure 2.4). Finally, **8OE3** was found to be nearly identical after superimposing it on **1MF5** and a RMSD of 0.103 Å was calculated.

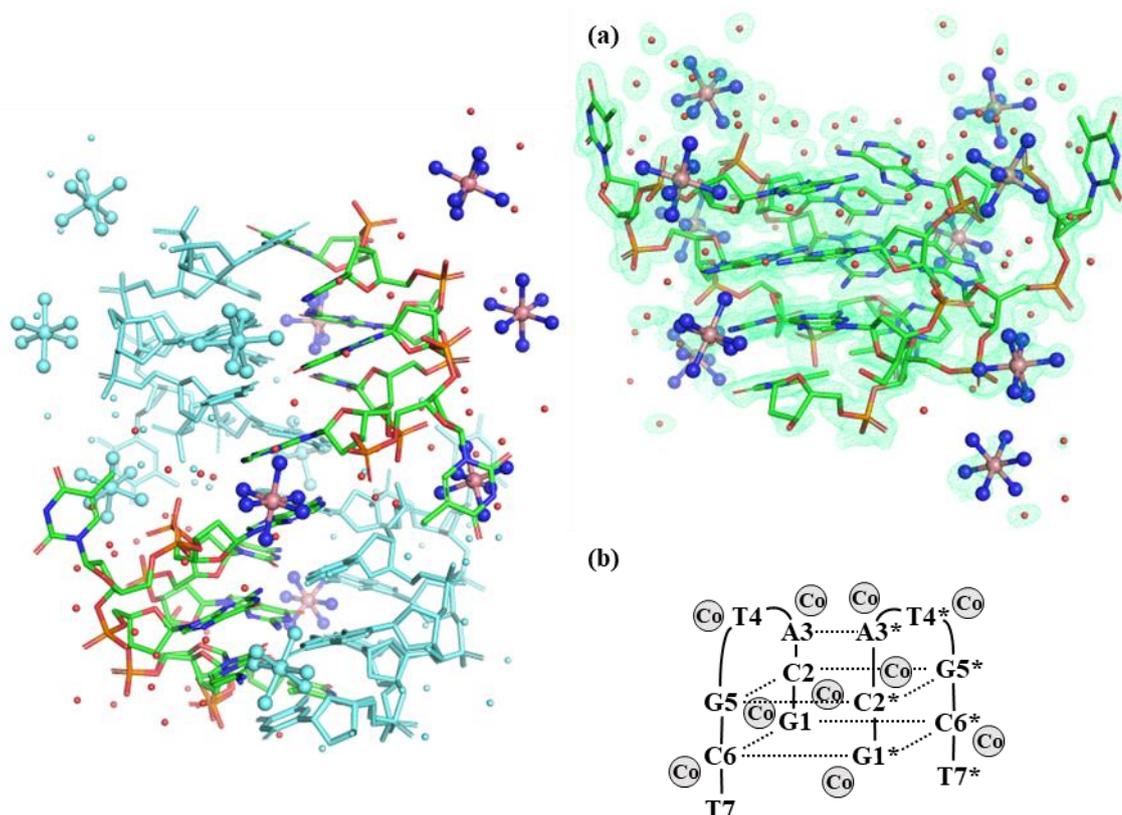


Figure 2.4: Asymmetric unit of **8OE3**. The two strands each form a non-canonical quadruplex with a symmetry mate. Carbon, oxygen, nitrogen, and phosphate atoms are shown as green, red, blue, and orange respectively. The symmetry mate is shown as cyan. a. Biological unit of **8OE3** with the $2mF_o-DF_c$ electron density map is contoured at the 1.0σ level (green). Carbon, oxygen, nitrogen, and phosphate atoms are shown as green, red, blue, and orange respectively. Hexamine cobalt (III) ions are shown as blue and pink spheres. Water molecules are shown as small red spheres. b. Schematic view of the biological unit of **8OE3** with the approximate positions of the hexamine cobalt (III) ions. Inter-strands and intra-strands interactions are shown as dashed lines.

8ASO and **8ASM** were obtained after soaking d(GCATGCT) crystals in nickel (II) chloride and cobalt (II) chloride respectively. The asymmetric unit was found to be two strands for a total of 14 nucleotides in both structures and the biological units form the same non-classical quadruplex structure as **1MF5** and **8OE3**. The structures pack with large solvent channels between the biomolecules, where the metal ions were found (Figure 2.5). However,

after calculating Matthews' coefficient in both **8ASO** and **8ASM**, the large solvent channels only accounted for 27.47 % and 26.87 % of the volume of the crystal respectively. **8ASO** and **8ASM** were superimposed on **8OE3** and low RMSD values (0.110 Å and 0.113 Å respectively) were calculated which suggests very little variability in the overall structures between before and after soaking. The internal structure has been found to be equal as well. The same numbering system was kept from **8OE3**. The nucleotides from strand A are labelled from G_{Ni}1 to T_{Ni}7 in the 5'-3' direction and the nucleotides from strand B are labelled from G_{Ni}8 to T_{Ni}14 in the 5'-3' direction for **8ASO**. The nucleotides from strand A are labelled from G_{Co}1 to T_{Co}7 in the 5'-3' direction and the nucleotides from strand B are labelled from G_{Co}8 to T_{Co}14 in the 5'-3' direction for **8ASM**. Nucleotides from symmetry related strands are designated with a *. Watson-Crick inter-strand bonds between G1*-C6 and G5*-C2 form the two G/C quadruplex formations, while Hoogsteen base pairs between A3-A3* and intra-strands bonds between G1-C6 and G5-C2 further stabilise the whole structure (Figure 2.4a). The thymine bases form single-base loops which stacks with other thymine bases of adjacent units.

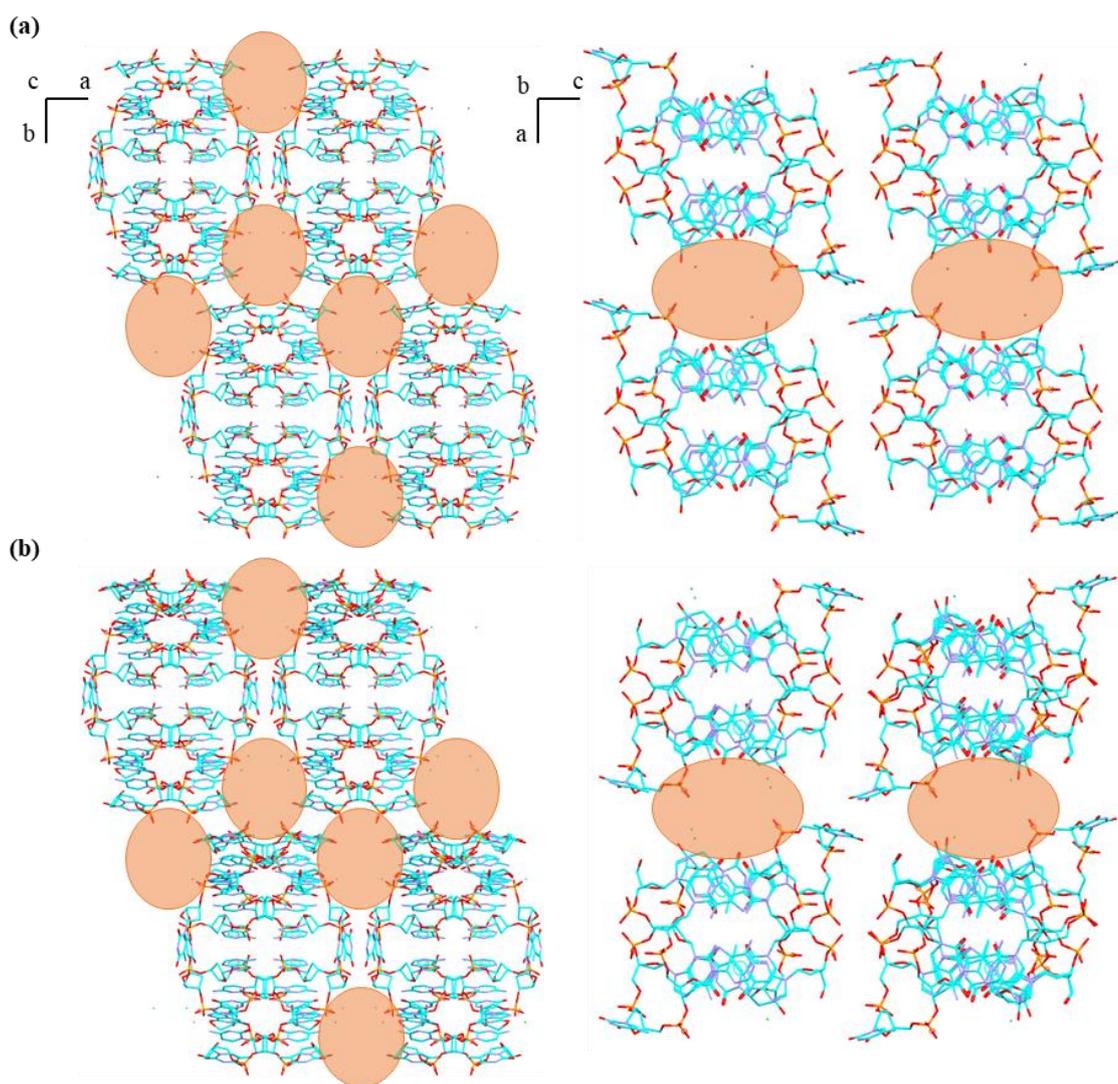


Figure 2.5: Packing diagrams for the structures **8ASM** (a) and **8ASO** (b) with views down the *c* (left) and *b* (right) axes. The quadruplexes are shown as sticks. Carbon, oxygen, nitrogen, and phosphate atoms are shown as cyan, red, blue, and orange respectively. Solvent channels are highlighted with orange circles. Metal ions are shown as spheres.

2.3.2. Comparison to results from cocrystallisation

Structures of the same sequence d(GCATGCT) obtained from crystals grown with nickel (II) chloride and cobalt (II) chloride, **1R2O**^a and **1QZL**^b respectively, have been deposited in the PDB in 2003. In both cases, the asymmetric unit consists of one strand for a total of seven nucleotides and one metal ion. However, the different growth methods led to a major structural difference between **8ASO/8ASM** and **1R2O/1QZL**, which radically changes the overall topology of the quadruplex. Unlike the two-stranded quadruplexes reported here

which formed in the absence of any metal ions, a non-classical four-stranded quadruplex is formed (Figure 2.6b). Like **8ASO** and **8ASM**, large solvent channels are present (Figure 2.6). However, 34.58 % and 34.30 % of the cell volume in **1R2O** and **1QZL** respectively was occupied by solvent. Compared to **8ASO** and **8ASM**, solvent volume increased by 7 %. In **1QZL**, the nucleotides are labelled from G_{Co-cocry}1 to T_{Co-cocry}7 in the 5'-3' direction. In **1R2O**, the nucleotides are labelled from G_{Ni-cocry}1 to T_{Ni-cocry}7 in the 5'-3' direction. Nucleotides from symmetry related strands are designated with *, ' and ". Made of two G/C quadruplexes stacked on top of each other, the internal structure is stabilised by Watson-Crick inter-strand bonds but only G1 and C2 are implicated in the G/C quadruplex (Figure 2.6b). Furthermore, the N7 position of the G1 bases involved in the G/C quadruplexes and the N7 position of the G5 bases present in the loops are bridged by a metal ion (Figure 2.6b & Figure 2.7). Finally, A3 and T4 form an internal loop and are stacked on top of the quadruplexes while C6 and T7 participate in the stacking.

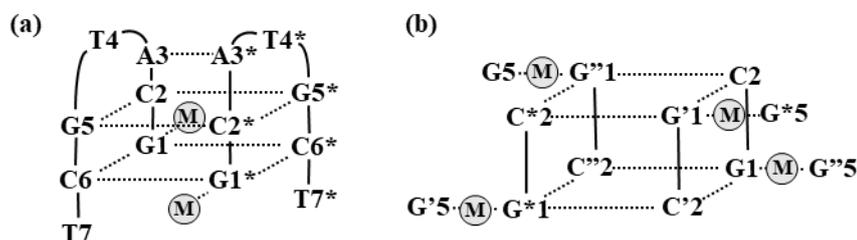


Figure 2.6: Comparison of 8ASO/8ASM and 1R2O/1QZL topologies. a. Schematic view of the biological unit of 8ASO/8ASM with the positions of the metal ions (M). b. Schematic view of the biological unit of 1R2O/1QZL with the positions of the metal ions (M). For clarity, A3, T4, C6 and T7 were not shown. Inter-strands interactions are shown as dashed lines.

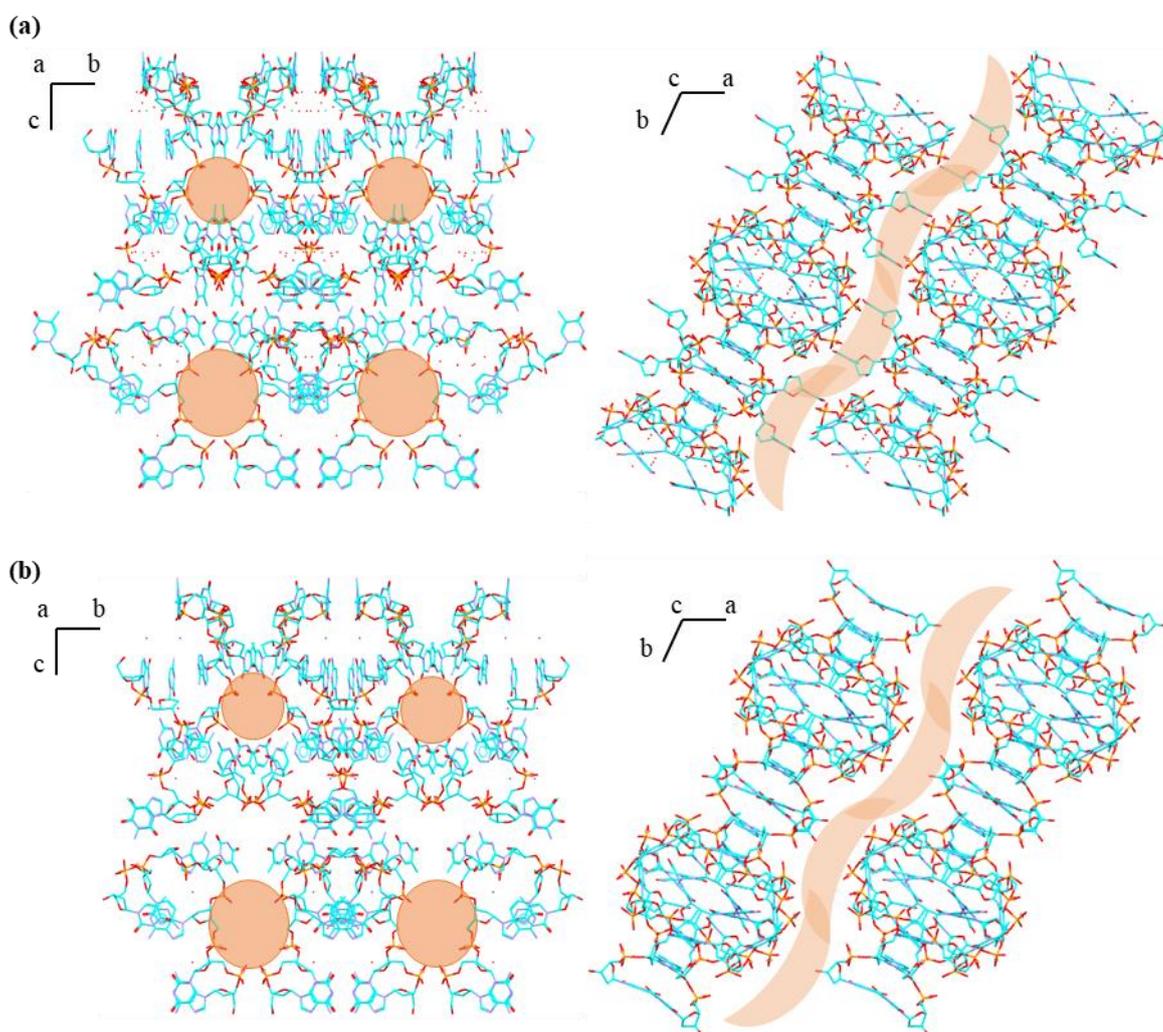


Figure 2.7: Packing diagrams for the structures 1R2O (a) and 1QZL (b) with views down the a (left) and c (right) axes. The quadruplexes are shown as sticks. Carbon, oxygen, nitrogen, and phosphate atoms are shown as cyan, red, blue, and orange respectively. Solvent channels are highlighted with orange circles. Metal ions and water molecules are shown as spheres.

2.3.3. Transition metal coordination analysis and solvent channel availability

In **8ASO** and **8ASM**, metal binding sites were found at the N7 positions of guanines. The Ni^{2+} ions were found directly coordinated to the N7 position of $\text{G}_{\text{Ni}1}$ and $\text{G}_{\text{Ni}7}$. Similarly, the Co^{2+} ions were found directly coordinated to the N7 position of $\text{G}_{\text{Co}1}$ and $\text{G}_{\text{Co}7}$ (Figure 2.8). Three positions were found for nickel ions, two of which were considered to be the same atom, split between two positions. After occupancy refinement, both positions of the nickel ion bound to $\text{G}_{\text{Ni}1}$ (Ni^{2+} (1a) and Ni^{2+} (1b)) were attributed an occupancy of 41 % and 59 %, while the nickel ion bound to $\text{G}_{\text{Ni}7}$ (Ni^{2+} (2)) was attributed an occupancy of 78 %. Equivalently, Co^{2+} (1)

and Co^{2+} (2) were found to be directly coordinated to $\text{G}_{\text{Co}1}$ (N7) and $\text{G}_{\text{Co}7}$ (N7) with a partial occupancy of 89 % and 68 % respectively. A hydration shell was expected and observed in both structures. All distances between metal ions and its close-by atoms are reported in Table 2.3.

Equivalently to **8ASO** and **8ASM**, the metal binding sites were also found at the N7 position of guanines. However, as described previously, the metal ions were found to bridge two guanines, G1 and G5 (Figure 2.8). Both the nickel ion (Ni^{2+} (3)) bound to $\text{G}_{\text{Ni-cocry}1}$ and the cobalt ion (Co^{2+} (3)) bound to $\text{G}_{\text{Co-cocry}1}$ were attributed an occupancy of 100 %. Metal ions parameters from **1R2O** and **1QZL** were also summarised in Table 2.3.

Table 2.3: Metals ions parameters from 8ASO, 8ASM, 1R2O and 1QZL (distance in Å).

PDB ID	Metal ion	Occupancy	Distance to guanine N7	Distance to water molecules
8ASO	Ni^{2+} (1a)	0.41	2.66	2.06; 2.05; 1.98; 2.25
	Ni^{2+} (1b)	0.59	1.93	-
	Ni^{2+} (2)	0.78	2.215 (A); 1.85 (B) ⁱ	2.57; 1.91; 2.27
8ASM	Co^{2+} (1)	0.89	2.54	1.94, 2.54
	Co^{2+} (2)	0.68	2.30	2.27
1R2O	Ni^{2+} (3)	1.00	2.25	2.15; 1.82; 2.14; 2.06
1QZL	Co^{2+} (3)	1.00	2.51	-

i – $\text{G}_{\text{Ni}1}$ has two conformations, A and B. The values represent the distances between Ni^{2+} (2) and conformation A and B of this guanine.

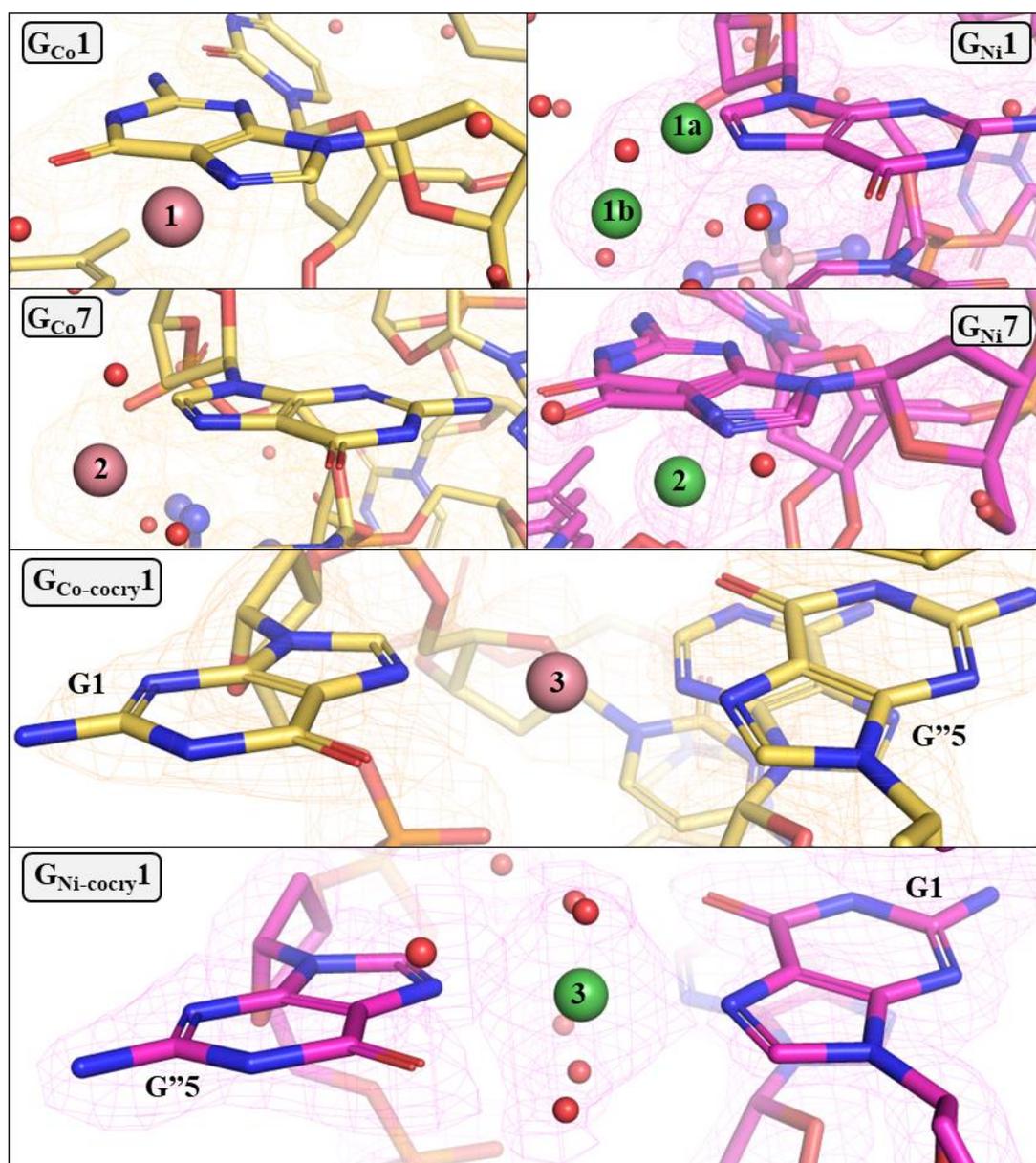


Figure 2.8: Comparison between all bound guanines present in the asymmetric unit of 8ASO, 8ASM, 1R2O and 1QZL. Carbon atoms are shown as yellow and magenta in the structures with cobalt and nickel ions respectively. Oxygen, nitrogen, and phosphate atoms are shown as red, blue, and orange respectively. Cobalt ions are shown as pink spheres. Nickel ions are shown as green spheres. Water molecules are shown as red spheres.

2.4. Discussion

The benefits of getting transition metals into a crystal system are many. Some transition metals such as Fe^{2+} and Cu^{2+} are crucial to enzymatic activity by catalysing oxidation-reduction reactions⁴¹ or others such as Ni^{2+} and Zn^{2+} were located in the active sites of enzymes⁴². Getting transition metals into the crystal can provide essential information to biological processes at an atomic level. In other circumstances, transition metals can be utilised to solve structures with the Single and Multi-wavelength Anomalous Dispersion (SAD/MAD) method^{18,19}, a method which uses transition metals as anomalous scatterers to determine structure factors and solve the “phase problem”. However, getting transition metals into a crystal can be quite challenging. To approach this problem, two methods can be employed: cocrystallisation and soaking, and each method comes with their advantages and disadvantages depending on which question the presence of metals would answer. In this thesis, preferred metal binding sites are studied as part of an investigation into the correlation between where oxidative DNA damage occurs and where a transition metal will bind on the DNA strand. Consequently, choosing carefully which method to use to get transition metals into a DNA crystal is crucial.

As can be observed from the two novel soaked structures described here, and the cocrystallised equivalents reported from the PDB, two types of non-canonical quadruplexes can form after introduction of transition metals to the DNA system. After soaking, nickel and cobalt ions were only able to bind to the guanine bases and were found in the solvent channels, a result which was expected, as previously, transition-metal ions such as Co^{2+} and Ni^{2+} have been reported to interact in the same way with DNA and to bind preferentially to the N7 position of guanines^{43–45}. Soaking of transition metals into protein or DNA crystals has been commonly described to identify transition-metal binding sites in crystal structures and has been the preferred method when cocrystallisation is unlikely to give results or when imposed time constraints require a much faster method⁴⁶. Moreover, soaking a crystal system which has already been described and having prior information on potential transition-metal binding sites and their availability from solvent channels could help characterise the system and understand binding specificity.

When crystals of this system are grown using cocrystallisation, Ni²⁺ and Co²⁺ ions are found to help stabilise the structure through the formation of intra-strand bridges between guanines. Co²⁺ and Ni²⁺ ions are important to the structural integrity of the molecular arrangement observed for structure **1R2O** and **1QZL**. And as reported previously, cocrystallisation of proteins or DNA with transition metals is more likely give the thermodynamically favoured structure⁴⁷. For example, the susceptibility of DNA to conformational changes induced by transition metals illustrates how transition metals promote thermodynamically favoured structures, which could indicate DNA might not necessarily crystallise as a B-DNA duplex^{48,49}. However, as proteins are not as flexible as DNA, transition metals might not promote as many thermodynamically favoured conformational changes, yet some proteins might display some flexibility to accommodate several metal ions in the active sites⁵⁰. Furthermore, thermodynamics have been described as an influence on where some of the metals bind in protein crystal structures and metal positions in protein structures might not reflect how metals are bound to the proteins in the natural environment⁴⁷. And while transition metals have been described as favouring negatively charged and polar amino acids, to predict where preferred transition metal binding sites are might be difficult in proteins⁵¹. By contrast, to predict where transition metals bind in DNA could prove easier as the bases and the groove are available most of the time, yet the structure after conformational changes might not be as predictable.

In this work, the DNA sequence d(GCATGCT) has been used to compare the two methods and determine which one would be the most appropriate to obtain ordered transition metals that could react with hydrogen peroxide molecules in the solvent channels. **8ASO** and **8ASM**, the two soaked structures, have been displaying similar transition-metal binding preferences comparable to what can be seen in other DNA systems soaked in transition metals. For example, crystals of the Z-DNA forming sequences d(CGCGCG) and d(CGCGTG) have been soaked in both Co²⁺ and Cu²⁺⁴³ and the authors showed that in the Z-DNA form, steric clashes caused between the hydrated metals bound to the N7 position of guanines and the previous base would influence the bound location. However, in **8ASO** and **8ASM**, the first shell of interaction, or the molecules and atoms directly coordinated to the metal ion, has only been partially observed. Metal binding selectivity is highly influenced by the first shell and second shell of interaction. However, the presence of a partial hydration spheres does not allow to

conclude if the position is favoured by a network of water molecules or if clashes with the Co^{2+} and Ni^{2+} ions could explain why they bind in priority to the G1 position of the strand in a disordered manner while the G5 position is not selected. However, the structure of the d(CGCGCG) sequence **1D39**⁵² might contribute to a better understanding in preferred binding sites in the presence of two adjacent guanine bases in a structure ahead of metal soaking. In comparison, the difference in distance between the atoms (Co^{2+} and water) bound to the two adjacent guanine bases in **8ASO** and **8ASM** and the atoms (Cu^{2+}) bound to the two adjacent guanine bases in **1D39** appears to be greater than 1 Å (Figure 2.9). Moreover, the ionic radii for Co^{2+} and Ni^{2+} are reported as 70 pm, 3 pm less than the ionic radius for Cu^{2+} ⁵³, a small difference which might explain the difference in binding profiles. Yet, closer potential metal ion centres with higher ionic radius did not prevent the copper (II) ions from binding in **1D39**. The distance between two potential metal centres alone cannot help predicting where potential metal centres are. Nonetheless, as the G5 positions are available from the solvent channels, another reasonable explanation for the presence of metal-free guanines could be low concentration of transition-metal solutions used for soaking and a too-short soaking time²⁵. However, the d(GCATGCT) crystals used for soaking would dissolve quite rapidly at higher concentration or at longer soaking times, especially when soaked in a drop containing nickel (II) chloride.

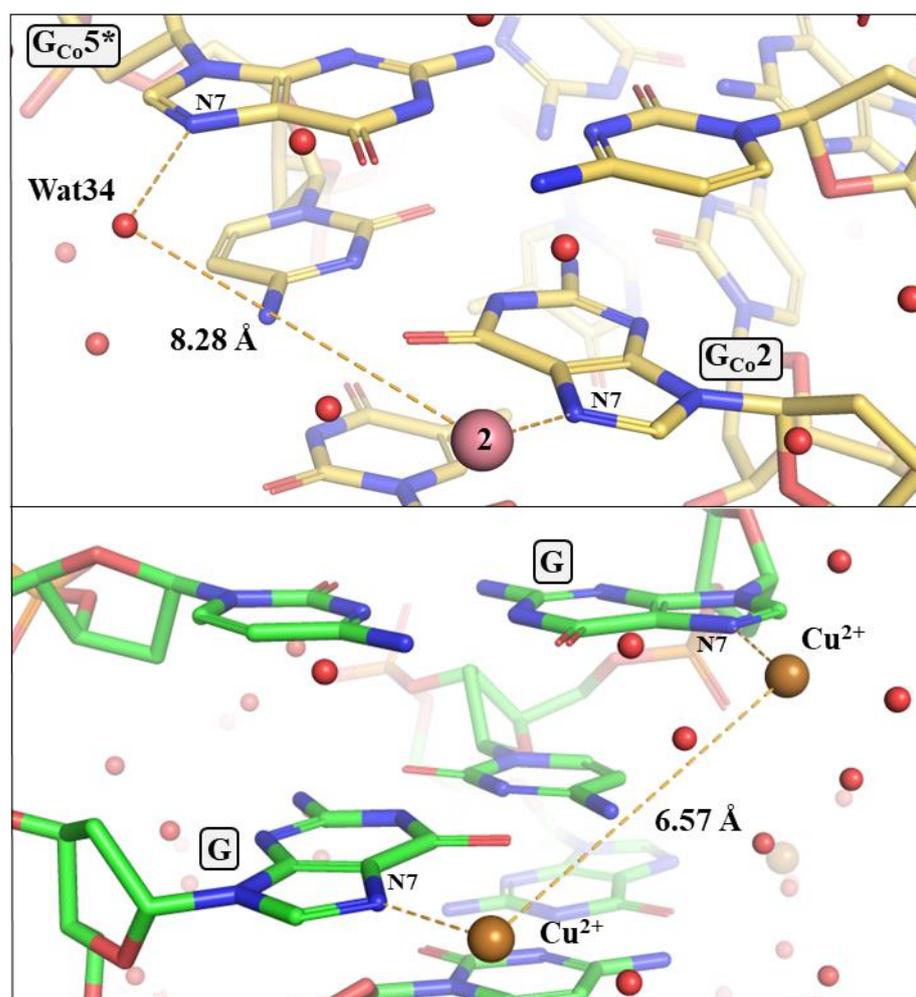


Figure 2.9: Comparison between two adjacent guanine bases and the distance between their metal centres in 8ASM (top) and 1D39⁵² (bottom). Carbon atoms are shown as yellow and green in 8ASM and 1D39 respectively. Oxygen, nitrogen, and phosphate atoms are shown as red, blue, and orange respectively. Cobalt ions are shown as pink spheres. Copper ions are shown as orange spheres. Water molecules are shown as red spheres.

On the contrary, guanine bridges were observed in both 1R2O and 1QZL after cocrystallisation, and all N7 positions were no longer available to the solvent. Moreover, the metal binding sites are deep within the structure and unavailable to the solvent, thus preventing any oxidation damages to occur. This reinforces the argument in favour of using soaking of metals instead of crystallisation in the examination of transition metal-DNA structures and DNA damage, especially since it is important the metal centres are available to the hydrogen peroxide molecules soaked into the crystal to successfully damage the DNA. Additionally, guanine-Ni²⁺-guanine bridges and guanine-Co²⁺-guanine bridges are particularly common in structures obtained from crystal grown in conditions containing these metal ions, where these

often occur among symmetry-related bases^{44,54–56}. However, the formation of cytosine/guanine quadruplexes stabilised by metal-ion bridges is unlikely *in vivo* as this would require duplex DNA in cells to be far more flexible than it already is. Additionally, extra-helical interactions between the transition-metals' hydration sphere and the phosphate backbone have also been reported to stabilise crystal packing in DNA duplex structures, interactions which might explain why the difference in method used influence crystal packing. Furthermore, since the cellular environment is not as concentrated in transition metal ions as crystallisation solutions are, with concentrations sometimes ten times higher than in physiological environments⁵⁷, there would be fewer free ions to form guanines bridges, and structures like **1R2O** and **1QZL** would be rarer. The structural flexibility of nucleic acids is important for many biological processes, from storage to replication and expression⁵⁸. And structures which might not be favoured in the cell environment due to low concentrations in transition metal could still be favoured during the cocrystallisation process and give structures with unavailable transition metals.

Separate studies have investigated the differences between soaking and cocrystallisation in their own systems. And although co-crystallization might be seen as more reliable to study other biological system like protein–drug interactions²⁶ and protein-ligands interactions²⁷, this may not be a technique which is suitable to obtain the nucleic acid structures containing transition metal centres needed to investigate damage. Additionally, the environment *in vivo* would be more likely to resemble a soaked environment, as the “correct” structure is able to form without transition-metals' intervention and assemble into an arrangement suitable for crystal growth.

2.5. References

1. Liu, J. *et al.* Metalloproteins Containing Cytochrome, Iron–Sulfur, or Copper Redox Centers. *Chem. Rev.* **114**, 4366–4469 (2014).
2. Shi, W. & Chance, M. R. Metalloproteomics: forward and reverse approaches in metalloprotein structural and functional characterization. *Curr. Opin. Chem. Biol.* **15**, 144–148 (2011).
3. Pyle, A. Metal ions in the structure and function of RNA. *JBIC J. Biol. Inorg. Chem.* **7**, 679–690 (2002).
4. Cassandri, M. *et al.* Zinc-finger proteins in health and disease. *Cell Death Discov.* **3**, 1–12 (2017).
5. Xi, K., Wang, F.-H., Xiong, G., Zhang, Z.-L. & Tan, Z.-J. Competitive Binding of Mg²⁺ and Na⁺ Ions to Nucleic Acids: From Helices to Tertiary Structures. *Biophys. J.* **114**, 1776–1790 (2018).
6. Auffinger, P., D’Ascenzo, L. & Ennifar, E. Sodium and Potassium Interactions with Nucleic Acids. in *The Alkali Metal Ions: Their Role for Life* 167–201 (Springer International Publishing, Cham, 2016).
7. Kolev, S. K. *et al.* Interaction of Na⁺, K⁺, Mg²⁺ and Ca²⁺ counter cations with RNA. *Metallomics* **10**, 659–678 (2018).
8. Gao, Y., Pink, M. & Smith, J. M. Alkali Metal Ions Dictate the Structure and Reactivity of an Iron(II) Imido Complex. *J. Am. Chem. Soc.* **144**, 1786–1794 (2022).
9. Pace, N. J. & Weerapana, E. Zinc-Binding Cysteines: Diverse Functions and Structural Motifs. *Biomolecules* **4**, 419–434 (2014).
10. Solomon, E. I. *et al.* Copper Active Sites in Biology. *Chem. Rev.* **114**, 3659–3853 (2014).
11. Kitatsuji, C. *et al.* Protein oxidation mediated by heme-induced active site conversion specific for heme-regulated transcription factor, iron response regulator. *Sci. Rep.* **6**, 18703 (2016).
12. Berman, H. M. *et al.* The Protein Data Bank. *Nucleic Acids Res.* **28**, 235–242 (2000).

13. Putignano, V., Rosato, A., Banci, L. & Andreini, C. MetalPDB in 2018: a database of metal sites in biological macromolecular structures. *Nucleic Acids Res.* **46**, 459–464 (2018).
14. Cowtan, K. Phase Problem in X-ray Crystallography, and Its Solution. in *eLS* (Wiley, 2003).
15. Taylor, G. L. Introduction to phasing. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 325–338 (2010).
16. Evans, P. & McCoy, A. An introduction to molecular replacement. *Acta Crystallogr. D Biol. Crystallogr.* **64**, 1–10 (2008).
17. Perutz, M. F. Isomorphous replacement and phase determination in non-centrosymmetric space groups. *Acta Crystallogr.* **9**, 867–873 (1956).
18. Hendrickson, W. A. & Teeter, M. M. Structure of the hydrophobic protein crambin determined directly from the anomalous scattering of sulphur. *Nature* **290**, 107–113 (1981).
19. Hendrickson, W. A. Determination of Macromolecular Structures from Anomalous Diffraction of Synchrotron Radiation. *Science* **254**, 51–58 (1991).
20. Morgan, B. & Lahav, O. The effect of pH on the kinetics of spontaneous Fe(II) oxidation by O₂ in aqueous solution – basic principles and a simple heuristic description. *Chemosphere* **68**, 2080–2084 (2007).
21. Santana-Casiano, J. M., González-Dávila, M. & Millero, F. J. Oxidation of Nanomolar Levels of Fe(II) with Oxygen in Natural Waters. *Environ. Sci. Technol.* **39**, 2073–2079 (2005).
22. Wedrychowski, A., Schmidt, W. N. & Hnilica, L. S. The *in vivo* cross-linking of proteins and DNA by heavy metals. *J. Biol. Chem.* **261**, 3370–3376 (1986).
23. Matthews, B. W. Solvent content of protein crystals. *J. Mol. Biol.* **33**, 491–497 (1968).
24. López-Jaramillo, F. J., Moraleda, A. B., González-Ramírez, L. A., Carazo, A. & García-Ruiz, J. M. Soaking: the effect of osmotic shock on tetragonal lysozyme crystals. *Acta Crystallogr. D Biol. Crystallogr.* **58**, 209–214 (2002).
25. Mizutani, R. *et al.* Spatiotemporal development of soaked protein crystal. *Sci. Rep.* **4**, 5731 (2014).

26. Wielen-Schmidt, B., Oebbeke, M., Ngo, K., Heine, A. & Klebe, G. Two Methods, One Goal: Structural Differences between Cocrystallization and Crystal Soaking to Discover Ligand Binding Poses. *ChemMedChem* **16**, 292–300 (2021).
27. Ehrmann, F. R. *et al.* Soaking suggests “alternative facts”: Only co-crystallization discloses major ligand-induced interface rearrangements of a homodimeric tRNA-binding protein indicating a novel mode-of-inhibition. *PLOS ONE* **12**, e0175723 (2017).
28. Thorpe, J. H., Teixeira, S. C. M., Gale, B. C. & Cardin, C. J. Crystal structure of the complementary quadruplex formed by d(GCATGCT) at atomic resolution. *Nucleic Acids Res.* **31**, 844–849 (2003).
29. Winter, G. xia2: an expert system for macromolecular crystallography data reduction. *J. Appl. Crystallogr.* **43**, 186–190 (2010).
30. Winter, G. *et al.* DIALS: implementation and evaluation of a new integration package. *Acta Crystallogr. Sect. Struct. Biol.* **74**, 85–97 (2018).
31. McCoy, A. J. *et al.* Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674 (2007).
32. Kantardjieff, K. A. & Rupp, B. Matthews coefficient probabilities: Improved estimates for unit cell contents of proteins, DNA, and protein–nucleic acid complex crystals. *Protein Sci.* **12**, 1865–1871 (2003).
33. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2126–2132 (2004).
34. Afonine, P. V. *et al.* Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr. D Biol. Crystallogr.* **68**, 352–367 (2012).
35. Kleywegt, G. J. & Brünger, A. T. Checking your imagination: applications of the free R value. *Structure* **4**, 897–904 (1996).
36. Read, R. J. *et al.* A New Generation of Crystallographic Validation Tools for the Protein Data Bank. *Structure* **19**, 1395–1412 (2011).
37. Winn, M. D. *et al.* Overview of the CCP4 suite and current developments. *Acta Crystallogr. D Biol. Crystallogr.* **67**, 235–242 (2011).

38. Schrödinger, LLC. The PyMOL Molecular Graphics System, Version 1.8. (2015).
39. Zheng, G., Lu, X.-J. & Olson, W. K. Web 3DNA—a web server for the analysis, reconstruction, and visualization of three-dimensional nucleic-acid structures. *Nucleic Acids Res.* **37**, W240–W246 (2009).
40. Macrae, C. F. *et al.* Mercury 4.0: from visualization to analysis, design and prediction. *J. Appl. Crystallogr.* **53**, 226–235 (2020).
41. MacPherson, I. S. & Murphy, M. E. P. Type-2 copper-containing enzymes. *Cell. Mol. Life Sci.* **64**, 2887–2899 (2007).
42. Ermler, U., Grabarse, W., Shima, S., Goubeaud, M. & Thauer, R. K. Active sites of transition-metal enzymes with a focus on nickel. *Curr. Opin. Struct. Biol.* **8**, 749–758 (1998).
43. Gao, Y.-G., Sriram, M. & Wang, A. H.-J. Crystallographic studies of metal ion - DNA interactions: different binding modes of cobalt(II), copper(II) and barium(II) to N7 of guanines in Z-DNA and a drug-DNA complex. *Nucleic Acids Res.* **21**, 4093–4101 (1993).
44. Abrescia, N. G., Huynh-Dinh, T. & Subirana, J. A. Nickel-guanine interactions in DNA: crystal structure of nickel-d[CGTGTACACG]₂. *JBIC J. Biol. Inorg. Chem.* **7**, 195–199 (2002).
45. Rohner, M., Medina-Molner, A. & Spingler, B. N,N,O and N,O,N Meridional cis Coordination of Two Guanines to Copper(II) by d(CGCGCG)₂. *Inorg. Chem.* **55**, 6130–6140 (2016).
46. Stewart, L., Clark, R. & Behnke, C. High-throughput crystallization and structure determination in drug discovery. *Drug Discov. Today* **7**, 187–196 (2002).
47. Barber-Zucker, S., Shaanan, B. & Zarivach, R. Transition metal binding selectivity in proteins and its correlation with the phylogenomic classification of the cation diffusion facilitator protein family. *Sci. Rep.* **7**, 16381 (2017).
48. Lee, J. S., Latimer, L. J. P. Reid, R. S. A cooperative conformational change in duplex DNA induced by Zn²⁺ and other divalent metal ions. *Biochem. Cell Biol.* **71**, 162–168 (1993).
49. Thirugnanasambandam, A., Karthik, S., Artheswari, G. & Gautham, N. DNA polymorphism in crystals: three stable conformations for the decadeoxynucleotide d(GCATGCATGC). *Acta Crystallogr. Sect. Struct. Biol.* **72**, 780–788 (2016).

50. Yeoh, K. K. *et al.* Dual-action inhibitors of HIF prolyl hydroxylases that induce binding of a second iron ion. *Org. Biomol. Chem.* **11**, 732–745 (2013).
51. Rulišek, L. & Vondrášek, J. Coordination geometries of selected transition metal ions (Co^{2+} , Ni^{2+} , Cu^{2+} , Zn^{2+} , Cd^{2+} , and Hg^{2+}) in metalloproteins. *J. Inorg. Biochem.* **71**, 115–127 (1998).
52. Kagawa, T. F., Geierstanger, B. H., Wang, A. H. & Ho, P. S. Covalent modification of guanine bases in double-stranded DNA. The 1.2-Å Z-DNA structure of d(CGCGCG) in the presence of CuCl_2 . *J. Biol. Chem.* **266**, 20175–20184 (1991).
53. Shannon, R. D. Revised effective ionic radii and systematic studies of interatomic distances in halides and chalcogenides. *Acta Crystallogr. A* **32**, 751–767 (1976).
54. Valls, N., Usón, I., Gouyette, C. & Subirana, J. A. A Cubic Arrangement of DNA Double Helices Based on Nickel–Guanine Interactions. *J. Am. Chem. Soc.* **126**, 7812–7816 (2004).
55. Thirugnanasambandam, A., Karthik, S., Mandal, P. K. & Gautham, N. The novel double-folded structure of d(GCATGCATGC): a possible model for triplet-repeat sequences. *Acta Crystallogr. D Biol. Crystallogr.* **71**, 2119–2126 (2015).
56. Valls, N., Wright, G., Steiner, R. A., Murshudov, G. N. & Subirana, J. A. DNA variability in five crystal structures of d(CGCAATTGCG). *Acta Crystallogr. D Biol. Crystallogr.* **60**, 680–685 (2004).
57. Finney, L. A. & O’Halloran, T. V. Transition Metal Speciation in the Cell: Insights from the Chemistry of Metal Ion Receptors. *Science* **300**, 931–936 (2003).
58. Marin-Gonzalez, A., Vilhena, J. G., Perez, R. & Moreno-Herrero, F. A molecular view of DNA flexibility. *Q. Rev. Biophys.* **54**, e8 (2021).

Chapter 3 - Damage Following Fenton-like Copper Reaction Influences Copper (II) Ions Binding Preferences to DNA Structure

3.1. Introduction

Copper (Cu) is an essential transition metal involved in many biological processes and found primarily in the body bound to ceruloplasmin in the plasma¹. However, its free hydrated form, Cu (II), has been studied for its ability to damage cellular components as the result of its reaction with hydrogen peroxide, a by-product of cellular respiration. Studies of Cu (II) and Cu (I) reactions with hydrogen peroxide (H₂O₂) have been reported^{2,3} and multiple mechanisms on the formation of reactive oxygen species (ROS) have been proposed. The formation of free radicals after catalysis by copper has been suggested, with the formation of O₂⁻ following the reaction between Cu (II) with H₂O₂. The excess H₂O₂, in turn, reacts with the reaction product Cu (I) to form HO[•]^{4,5}. However, recently, it has also been postulated that HO[•] is not as important in the system as it was thought as the reaction between Cu (I) and H₂O₂ did result instead in the oxidant Cu (III)⁶.

In humans, dysregulation of metal levels including iron or copper has been reported in neurodegenerative diseases, including Parkinson's and Alzheimer's disease, and oxidative stress and damage have been correlated to these changes in metal concentrations⁷. The oxidative stress marker product of nucleic acid damage, 8-oxo-2'-deoxyguanine (oxoG), has been used to assess oxidative stress levels in cells⁸⁻¹¹. Several studies have reported increased formation of base lesions compared to strand-breaks after Cu-mediated DNA damage. Cu (II) and Cu (I) binding of DNA has been shown to occur primarily at the N7 position of guanines, a position which has been described to help "deliver hydroxyl radical via binding" to form the intermediate radical to oxoG formation, 8-HOdG¹². Moreover, the more successive guanines there are in a tract, the more prone to oxidation the guanines are due to π stacking, a phenomenon calculated by Saito *et al.*¹³ in their work. However, it has remained unclear if bound metal positions correlate with base damage, especially since little work has been done to look at metal binding preferences in the presence of neighbouring or successive guanine bases.

Copper centres have been previously investigated using X-ray crystallography. Seven structures containing cupric ions have been deposited on the Protein Data Base to investigate binding preferences of copper (II) ions. However, while Z-DNA is heavily represented, with six out of seven structures reported as a Z-DNA¹⁴⁻¹⁶, binding preferences to many canonical and non-canonical nucleic acid structures have not been reviewed at an atomic level. Furthermore, techniques including Circular Dichroism and Infrared Spectroscopy have been used previously to look copper interactions with chromatin and have shown copper (II) ions to alter the B-DNA conformation, an outcome associated with potential copper induced damage and alteration of normal processes including DNA replication and transcription^{17,18}. Diversifying structures would however prove useful to better understand oxidative damage preferences and identify damage hot spots in a three-dimensional environment.

d(GCATGCT) has been shown to form a non-canonical quadruplex and as discussed in Chapter 2, to be a reliable system to soak in transition metals. This structure could prove to be a good system to investigate the binding properties of copper in the presence of two neighbouring guanines and the effects of one of the two guanines getting damaged after oxidative damage. After successfully soaking d(GCATGCT) crystals with nickel and cobalt, the same method was applied with copper (II) chloride. The result was compared to the structures **8ASO** and **8ASM** reported previously in Chapter 2, as well as to two novel structures of the same sequence containing the major oxidation product 8-oxo-7,8-dihydroguanine.

3.2. Material and methods

3.2.1. DNA preparation

The oligonucleotides d(GCATGCT) and d((oxoG)CATGCT) were purchased from Eurogentec as solid purified by RP-HPLC. DNA concentration was determined from the value of the absorbance at $\lambda = 260$ nm, using the molar extinction coefficient provided by the manufacturer. Both sequences were prepared to 1 mM and were annealed in a 40 mM sodium cacodylate trihydrate pH 5.7 buffer at 90°C for two minutes then allowed to cool slowly to room temperature.

3.2.2. Crystallisation of d(GCATGCT) and d((oxoG)CATGCT) and soaking with copper (II) chloride

Single crystals of d(GCATGCT) and d((oxoG)CATGCT) to be used for soaking with copper were grown using the same protocol as described in Chapter 2. The drops contained 2 μ L of 450 μ M oligonucleotides and 2 μ L of the crystallisation solution containing 2.20 % (v/v) MPD, 8.80 mM sodium cacodylate trihydrate pH 5.7, 17.80 mM hexamine cobalt (III), 13.30 mM sodium chloride and 355 mM potassium chloride. The drops were equilibrated against 500 μ l of 35 % (v/v) MPD. The crystals grew within three months and were soaked in a drop containing 2 μ L of the crystallisation conditions and 2 μ L of 25 mM copper (II) chloride for 5 days and 24 hours, respectively.

3.2.3. Data collection and data analysis for a crystal of d(GCATGCT) with copper (II) chloride and a crystal of native d((oxoG)CATGCT)

Data from the d(GCATGCT) crystal soaked in copper (II) chloride and the d((oxoG)CATGCT) crystal were collected from a flash-cooled single crystal of approximate dimensions 200 x 100 x 100 μ m and 150 x 100 x 100 μ m respectively, at 100 K. The data were collected using the in-house XtaLAB Synergy-S diffractometer using a PhotonJet-S X-ray source at a wavelength of 1.5406 Å with a HyPix-6000HE detector. The data for the copper-soaked crystal were integrated and scaled using the CrysAlisPRO software, giving 9598 unique reflections. The structure was solved by molecular replacement using Phaser¹⁹ and the structure

from the PDB **1MF5**. The model was updated using Coot²⁰ and refined using phenix.refine²¹. To confirm that the model was satisfactory and that the observed data and the calculated data from the model were in agreement, the R_{free} and the R_{work} were calculated after each round of refinement. R-factors can be made arbitrarily low by overfitting the data during refinement. Consequently, five percent of the reflections were reserved for the R_{free} set and not used during refinement. The R_{free} and R_{work} observed after the final round of refinement for this model were 0.1819 and 0.1349 respectively. These values can be compared relatively to X-ray structures of similar resolution after deposition and validation in the Protein Data Bank. The data, and final coordinates, were deposited in the Protein Data Bank with ID **8BAE**.

The data for the oxidised structure were integrated and scaled using the CrysAlisPRO software, giving 4151 unique reflections. The structure was solved by molecular replacement using Phaser and the structure from the PDB **1MF5**²². The model was updated using Coot. Both oxoguanines were refined using the ligand parameters from entry 8OG using eLBOW²³. Refinement was done using phenix.refine and give a final R_{work} of 0.1861 and R_{free} of 0.1969. Five percent of reflections were reserved for the R_{free} set. The data, and final coordinates, were deposited in the Protein Data Bank with ID **8BAF**. The data collection and refinement statistics for **8BAE** and **8BAF** can be found in Table 3.1.

Data from the d((oxoG)CATGCT) crystal soaked in copper (II) chloride were collected from a flash-cooled single crystal of approximate dimensions 150 x 100 x 100 μm , at 100 K with 0.004 s frames on beam line I03 at Diamond Light Source Ltd using radiation with a wavelength of 0.9762 Å. 3600 frames were collected with an oscillation angle of 0.1° to give 360° of data. The data were integrated and scaled using xia2²⁴ with XDS²⁵ giving 11191 unique reflections. The (oxoGCATGCT)-Cu structure was solved by molecular replacement using Phaser and the unsoaked structure presented in this chapter. The model was updated using Coot and refined using phenix.refine to give a final R_{work} of 0.1518 and R_{free} of 0.1724. Five percent of reflections were reserved for the R_{free} set. The data, and final coordinates, were deposited in the Protein Data Bank with ID **8BAG**. The data collection and refinement statistics for **8BAG** can be found in Table 3.1.

Table 3.1: Data collection and refinement statistics for 8BAE, 8BAF and 8BAG. Outer shell statistics shown in parentheses.

PDB ID	8BAE	8BAF	8BAG
Space group	C222 ₁	C222 ₁	C222 ₁
Cell dimensions			
<i>a</i> , <i>b</i> , <i>c</i> , (Å)	21.95, 58.87, 45.60	22.03, 58.85, 45.49	21.99 58.87 45.46
Resolution, (Å)	20.57-1.20 (1.22-1.20)	18.79-1.60 (1.64-1.60)	29.43-1.13 (1.15-1.13)
R _{merge}	0.099 (0.498)	0.077 (0.143)	0.070 (0.964)
R _{meas}	0.103 (0.538)	0.086 (0.175)	0.073 (1.022)
R _{pim}	0.028 (0.198)	0.036 (0.100)	0.021 (0.329)
Total number of observations	108640	19619	139516
Total number of unique observations	9598	4151	11191
CC _{1/2}	0.999 (0.873)	0.996 (0.958)	0.999 (0.756)
I/σI	14.7 (3.6)	13.6 (6.0)	15.3 (1.9)
Completeness, %	99.8 (97.0)	99.8 (99.3)	97.9 (89.2)
Multiplicity	11.3 (7.1)	4.7 (2.9)	12.5 (9.2)
Refinement			
Resolution,	20.57 - 1.20 (1.243 - 1.20)	18.792 - 1.600 (1.657 - 1.600)	29.435 - 1.130 (1.170 - 1.130)
No. Reflections	9566	4128	11162
R _{work} /R _{free}	0.1349/0.1819	0.1861/0.1969	0.1518/0.1724
rmsd			
Bond lengths, (Å)	0.006	0.022	0.017
Bond angles, (°)	0.87	2.04	1.76

3.2.4. Structural analysis

All figures were created using PyMOL²⁶ unless specified otherwise. Dinucleotide step parameters were calculated using Web 3DNA 2.0²⁷ are available in the Supplementary Information. CCP4 maps were generated from the final mtz file, including the anomalous map when available, using the phenix.mtz2map program. Maps were visualised in PyMOL and displayed using the mesh function. *2mFo-DFc* maps were contoured at 1σ unless specified otherwise. Anomalous maps were contoured at 3σ unless specified otherwise. The surface maps generated were visualised in PyMOL and displayed using the mesh function at level 2.0. Packing diagrams were generated using Mercury²⁸.

3.3. Results

3.3.1. oxoG substitution in a non-classical quadruplex structure

Here is reported the structure (PDB ID: **8BAF**) acquired after substituting the guanine in position 1 of the d(GCATGCT) sequence by its oxidisation product 8-oxoguanine. The structure was found to be two strands, of which each form a non-canonical quadruplex with a symmetry mate, for a total of 14 nucleotides (Figure 3.1). The biological unit is comprised of two strands forming a non-classical quadruplex structure. The structure was superimposed to the native structure reported in Chapter 2, **8OE3**, and **8BAF** was found to be identical to **8OE3** after a RMSD of 0.137 Å was calculated. The nucleotides are labelled from oxoG1 to T7 in the 5'-3' direction on strand A and from oxoG8 to T14 in the 5'-3' direction on strand B. Nucleotides from symmetry related strands are designated with a *. The Watson-Crick inter-strand bonds between oxoG1*-C6, C6*-oxoG1, C2*-G5 and G5*-C2, which form the two G/C quadruplex formations, and the Hoogsteen base pair between A3-A3* reported in **8BAF** are also reported in **8OE3** (Figure 3.1a). Additionally, no alternative conformations for G1*, C6* and T7* were found in **8BAF**, unlike in **8OE3**. Four hexamine cobalt residues were also reported in **8BAF**'s asymmetric unit compared to the six in **8OE3**. Additionally, crystal packing analysis showed no difference in the solvent channels compared to the ones found in the undamaged structure (Figure 3.2).

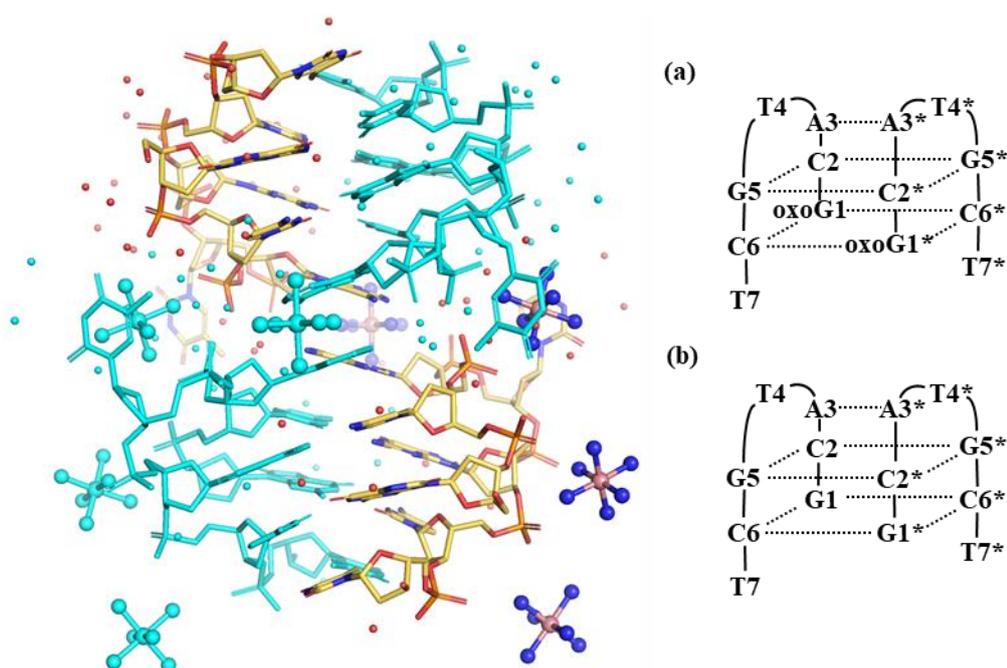


Figure 3.1: Asymmetric unit of **8BAF**. The two strands each form a non-canonical quadruplex with a symmetry mate. Carbon, oxygen, nitrogen, and phosphate atoms are shown as yellow, red, blue, and orange respectively. The symmetry mate is shown as cyan. Hexamine cobalt (III) ions are shown as blue and pink spheres. Water molecules are shown as small red spheres. a. Schematic view of the biological unit of **8BAF**. b. As comparison, the schematic view of the biological unit of **8OE3**. Inter-strands interactions are shown as dashed lines.

Locally, both oxoguanine bases participate in the quadruplexes and the reported local base-pair parameters from the damaged structure show little difference with the parameters from the native structure. Details of the local base-pair parameters can be found in the Supplementary Information. Additionally, the sugar puckers in **8BAF** have been classified as C2'-endo for both oxoG1 and oxoG8 and were found not to differ from the sugar puckers reported for G1 and G8 in **8OE3**. (Figure 3.3) Comparison of sugar puckering can be found in Table 3.2.

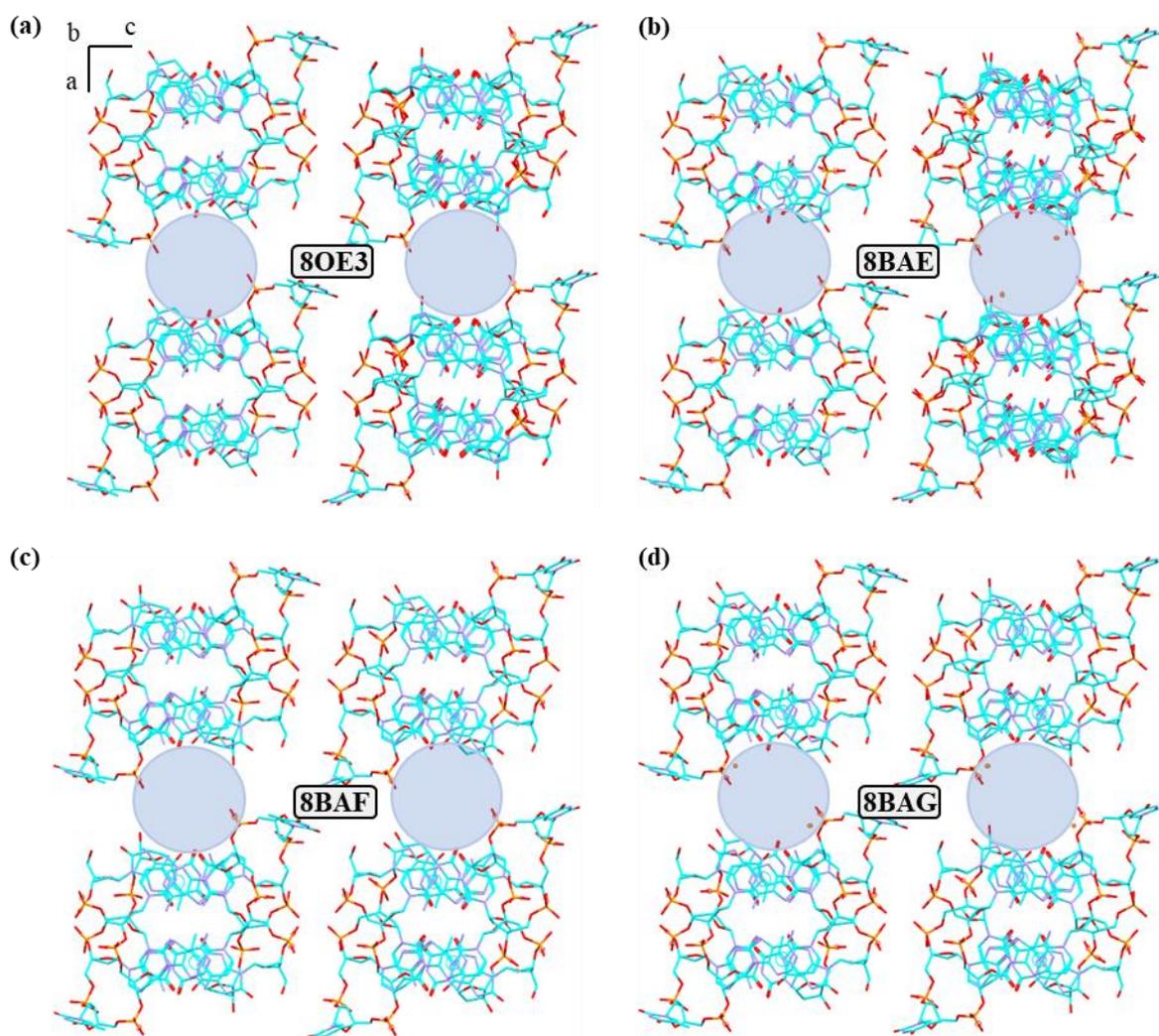


Figure 3.2: Packing diagrams for the structures 8OE3 (a), 8BAE (b), 8BAF (c) and 8BAG (d) with views down the b axis. The quadruplexes are shown as sticks. Carbon, oxygen, nitrogen, and phosphate atoms are shown as cyan, red, blue, and orange respectively. Solvent channels are highlighted with blue circles. The copper (II) ions are shown as spheres.

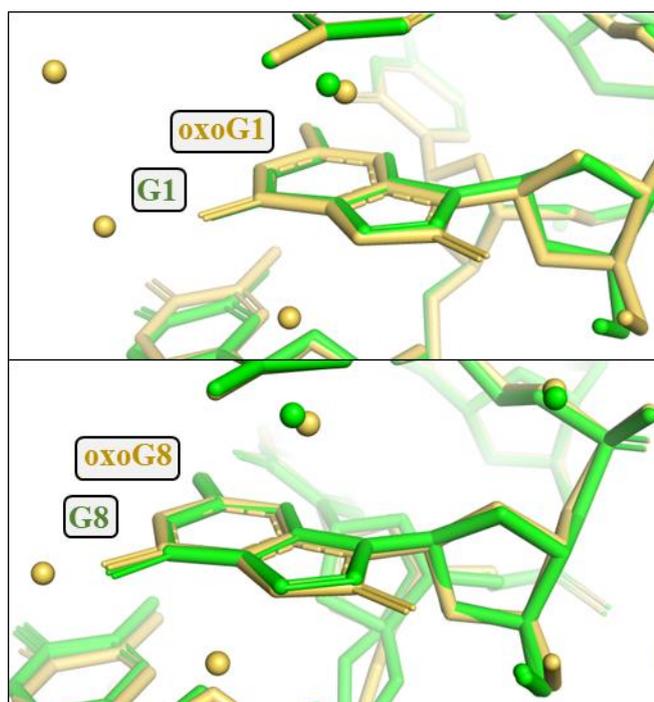


Figure 3.3: Close-up of the superposition between oxoG1 and G1 (top) and oxoG8 and G8 (bottom) from the structures 8BAF and 8OE3 respectively. 8BAF and 8OE3 are shown as yellow and green sticks respectively. Water molecules are shown as spheres.

Table 3.2: Sugar pucker pseudorotation angles and sugar pucker type for the oligonucleotide d(GCATGCT) and d(8oxoGCATGCT) in 8OE3 and 8BAF respectively.

		8OE3		8BAF		
Base		Phase angle of pseudorotation (P), °	Sugar Pucker	Base	Phase angle of pseudorotation (P), °	Sugar Pucker
C H A I N A	5'-G	177.8	C2'-endo	5'-oxoG	163.4	C2'-endo
	C	117.7	C1'-exo	C	117.3	C1'-exo
	A	157.4	C2'-endo	A	155.9	C2'-endo
	T	170.3	C2'-endo	T	167.7	C2'-endo
	G	183.2	C3'-exo	G	180.6	C3'-exo
	C	124.2	C1'-exo	C	130.9	C1'-exo
	T-3'	11.0	C3'-endo	T-3'	184.5	C3'-exo
C H A I N B	5'-G	184.8	C3'-exo	5'-oxoG	172.3	C2'-endo
	C	134.5	C1'-exo	C	125.8	C1'-exo
	A	153.3	C2'-endo	A	152.3	C2'-endo
	T	156.7	C2'-endo	T	159.4	C2'-endo
	G	186.5	C3'-exo	G	181.2	C3'-exo
	C	123.1	C1'-exo	C	105.0	O4'-endo
	T-3'	167.6	C2'-endo	T-3'	356.5	C2'-exo

3.3.2. Copper (II) ions coordination analysis after soaking of native crystals

8BAE was obtained after soaking native crystals of d(GCATGCT) in copper (II) chloride. Like **8OE3**, the structure was found to be two strands, of which each form a non-canonical quadruplex with a symmetry mate, for a total of 14 nucleotides (Figure 3.4). The biological unit is comprised of two strands forming a non-classical quadruplex structure. The structure was also superimposed to the native structure **8OE3**, and **8BAE** was found to be nearly identical to **8OE3** after a RMSD of 0.106 Å was calculated, which indicated soaking has not induced any significant changes. The nucleotides are labelled from G1 to T7 in the 5'-3' direction on strand A and from G8 to T14 in the 5'-3' direction on strand B. Nucleotides from symmetry related strands are designated with a *. Equivalently, the Watson-Crick inter-strand bonds between G1*-C6, C6*-G1, C2*-G5 and G5*-C2, which form the two G/C quadruplex formations, and the Hoogsteen base pair between A3-A3* reported in **8BAE** were reported in **8OE3** (Figure 3.4a). Additionally, alternative conformations at the G1* 5'-OH group, along with disorder at the C6*-T7* phosphate linkage, were reported with 60/40 % and 62/38 % occupancies respectively. Four hexamine cobalt molecules were reported in **8BAF**'s asymmetric unit compared to the six in **8OE3**. Additionally, analysis of the crystal packing showed no difference either between the solvent channels in **8BAE** and **8OE3** (Figure 3.2).

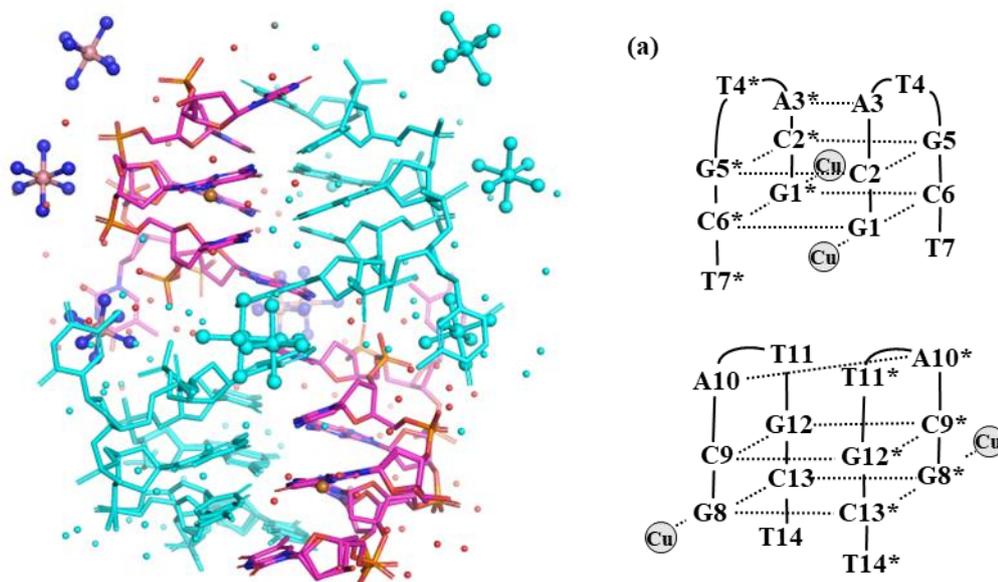


Figure 3.4: Asymmetric unit of 8BAE. The two strands each form a non-canonical quadruplex with a symmetry mate. Carbon, oxygen, nitrogen, and phosphate atoms are shown as magenta, red, blue, and orange respectively. The symmetry mate is shown as cyan. Hexamine cobalt (III) ions are shown as blue and pink spheres. Water molecules are shown as small red spheres. a. Schematic view of the asymmetric unit of 8BAE and the symmetry mate for each strand. Nucleotides from symmetry related strands are designated with a *. Inter-strands interactions are shown as dashed lines.

After soaking, the Cu^{2+} ions were found directly coordinated to the N7 position of G1 and G8 (Figure 3.5). After occupancy refinement, the copper ion bound to G1 ($Cu^{2+}(1)$) was attributed an occupancy of 66 %, while the copper ion bound to G8(A) and G8(B) ($Cu^{2+}(2)$) was attributed an occupancy of 22 %. Noticeably, no Cu^{2+} ions were found directly coordinated to G5 and G12 (Figure 3.5). A partial hydration shell was observed around both copper ions. Both metal ions parameters were reported in Table 3.3.

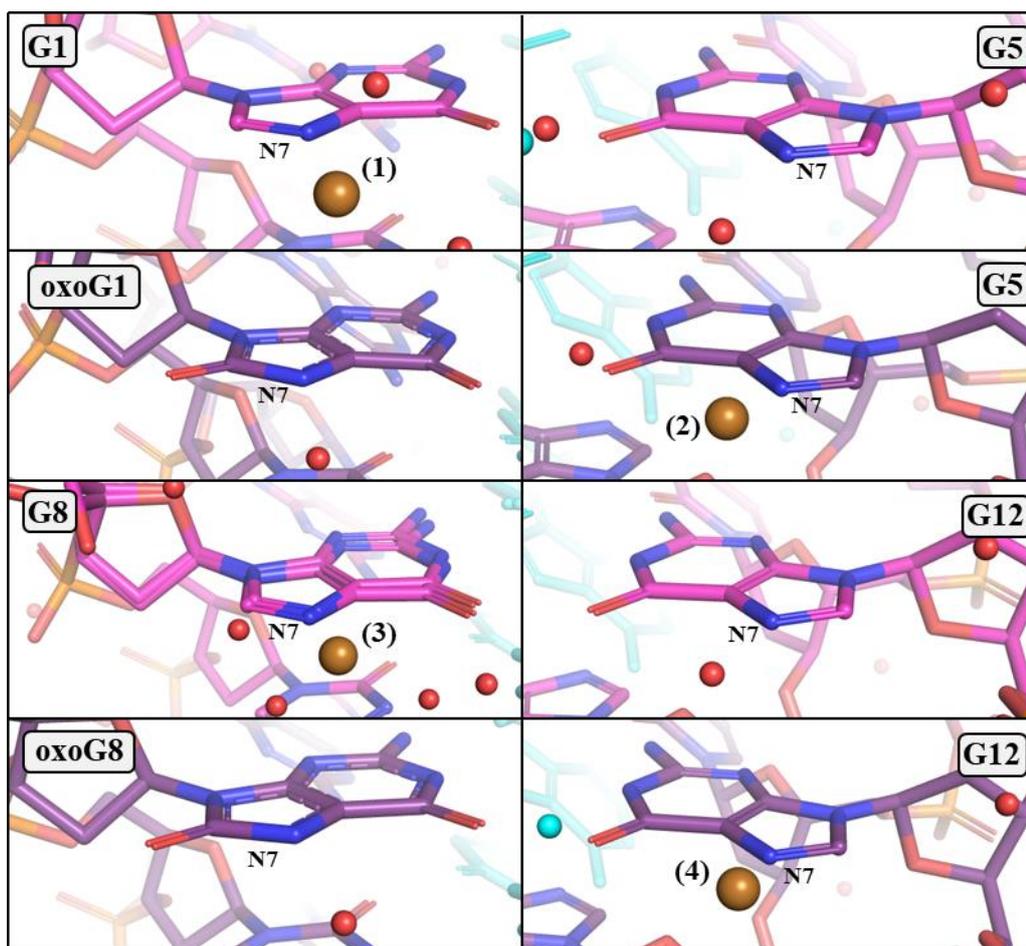


Figure 3.5: Comparison in copper (II) ions coordination between 8BAE (magenta) and 8BAG (purple). All the guanine bases present in the asymmetric units of the two structures are represented. Carbon atoms are shown as magenta and purple in 8BAE and 8BAG respectively. Oxygen, nitrogen, and phosphate atoms are shown as red, blue, and orange respectively. Copper ions are shown as orange spheres. Water molecules are shown as red spheres. Atoms from the symmetry mate are shown as cyan.

Table 3.3: Metals ions parameters from 8BAE and 8BAG (distance in Å)

PDB ID	Metal	Occupancy	Distance to guanine N7	Distance to water molecules
8BAE	Cu ²⁺ (1)	0.66	2.40	2.08; 2.02; 3.00
	Cu ²⁺ (2)	0.22	2.02 (A); 1.87 (B)	1.54; 1.94; 1.86; 2.84
8BAG	Cu ²⁺ (3)	0.39	2.19	2.50; 3.23
	Cu ²⁺ (4)	0.34	2.18	2.97; 3.12

Like **8BAE**, **8BAG** was obtained after soaking native crystals of d(oxoGCATGCT) in copper (II) chloride. And like the other structures from this work, the structure was found to be two strands, of which each form a non-canonical quadruplex with a symmetry mate, for a total of 14 nucleotides (Figure 3.6). The biological unit is comprised of two strands forming a non-classical quadruplex structure. The structure was also superimposed on its native structure **8BAF**, and **8BAG** was also found to be nearly identical to **8BAF** after a RMSD of 0.106 Å was calculated, which indicated soaking has not induced any significant changes. The nucleotides are labelled from G1 to T7 in the 5'-3' direction on strand A and from G8 to T14 in the 5'-3' direction on strand B. Nucleotides from symmetry related strands are designated with a *. Equivalently, the Watson-Crick inter-strand bonds between G1*-C6, C6*-G1, C2*-G5 and G5*-C2, which form the two G/C quadruplex formations, and the Hoogsteen base pair between A3-A3* reported in **8BAG** were reported in **8OE3** (Figure 3.6a). Furthermore, like in **8BAF**, no disorder was reported. Additionally, no difference was found between the solvent channels found in **8BAG** and **8BAF** either. Crystal packing was consistent between all the three figures presented in this study (Figure 3.2).

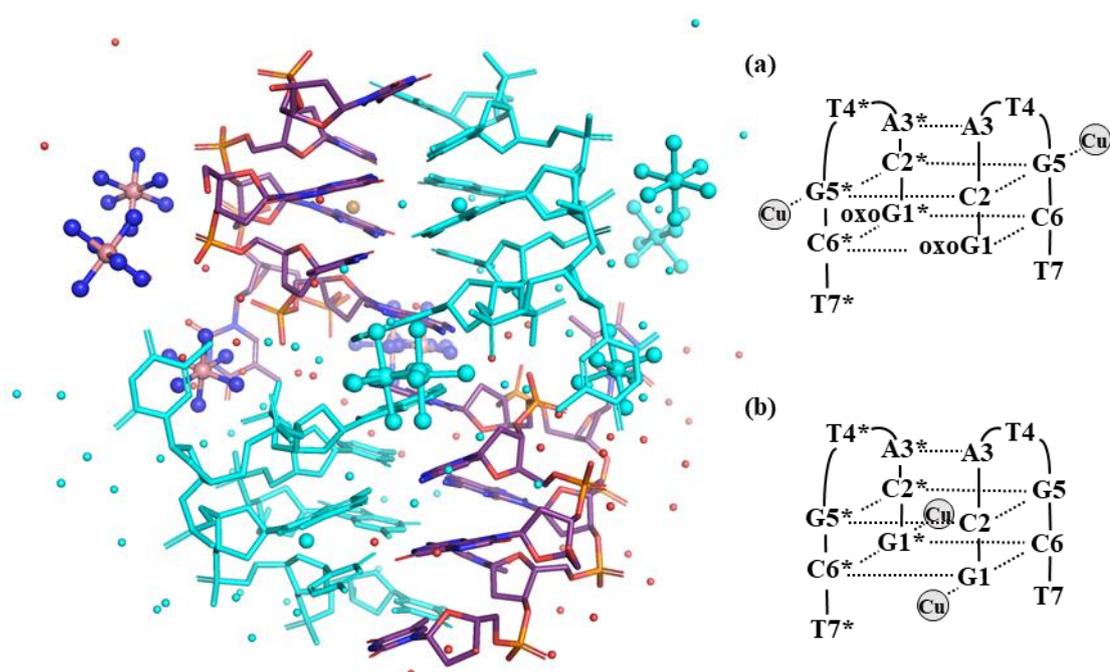


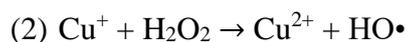
Figure 3.6: Asymmetric unit of 8BAG. The two strands each form a non-canonical quadruplex with a symmetry mate. Carbon, oxygen, nitrogen, and phosphate atoms are shown as purple, red, blue, and orange respectively. The symmetry mate is shown as cyan. Hexamine cobalt (III) ions are shown as blue and pink spheres. Water molecules are shown as small red spheres. a. Schematic view of the biological unit of 8BAG with the location of the copper (II) ions. The copper (II) ions are coordinated to the first guanine present in the strand. b. In comparison, schematic view of the biological unit of 8BAE with the location of the copper (II) ions. The copper (II) ions are coordinated to the fifth guanine present in the strand. Nucleotides from symmetry related strands are designated with a *. Inter-strands interactions are shown as dashed lines.

Comparatively to **8BAE**, the Cu^{2+} ions were found in **8BAG** after soaking. However, unlike **8BAE**, the Cu^{2+} ions were found directly coordinated to the N7 position of G5 and G12 (Figure 3.5). After occupancy refinement, the copper ion bound to G5 ($\text{Cu}^{2+}(3)$) was attributed an occupancy of 39 %, while the copper ion bound to G12 ($\text{Cu}^{2+}(4)$) was attributed an occupancy of 34 %. Both metal ions parameters were also reported in Table 3.2.

3.4. Discussion

3.4.1. Effects of oxidative damage on metal ions coordination

Cu-mediated oxidative damage of nucleic acids has been extensively studied and versions of this reaction have been used to investigate both the reactive species and the resulting damage in the presence of both copper and hydrogen peroxide. The reaction between hydrogen peroxide and Cu (II) produces Cu (I) and the reactive oxygen species $O_2^{\cdot-}$ (Scheme 3.1.(1)) and is known to be slow and the limiting step of the redox cycle³. Therefore, reducing agents such as ascorbic acid or hydroxylamine have been used to accelerate the reaction rate^{12,29}. Cu (I) reaction with H_2O_2 is also known to produce $HO\cdot$, a short-lived oxygen species which reacts at diffusion rates with nearby molecules, including nucleic acids³⁰ (Scheme 3.1.(2)).



Scheme 3.1: The reaction of copper with hydrogen peroxide

Hydroxyl radicals play a significant role in causing oxidative damage to nucleic acids and certain bases, and the proximity of the copper reaction centre when bound to the N7 position of guanines has been put forward as an explanation to base damage being more prevalent when copper is present¹².

Preferred binding sites for copper (II) ions in two systems forming non-canonical quadruplexes have been investigated in this study to determine possible damage hot spots in the event a guanine base is oxidised after the reaction between a copper (II) ion bound to the N7 positions of the guanine base and a hydrogen peroxide molecule. Copper coordination was determined at the N7 position of the guanine present at the 5' end of both strands, in a similar way as nickel and cobalt in **8ASM** and **8ASO** as discussed in Chapter 2. This suggests the same unavailability of the N7 position of the second guanine (G5 and G12) for transition metals to coordinate, possibly because of a metal centre more favourable at the G1 and G8 positions, as copper binding can be influenced by both nucleophilicity of the binding site and steric effects

around it³¹. However, as both G1 and G5 are available from the solvent channels, these positions should not be unavailable to copper (II) coordination when only potential binding sites availabilities are considered. Furthermore, the distance between G1 and G5* is close to the distance which has been observed between two neighbouring guanines in the Z-DNA forming sequence d(CGCGCG)¹⁴, so the proximity of two potential binding sites for copper (II) is unlikely to be determinant in binding preferences here. However, the absence of a complete solvation sphere around all the copper (II) ions found in both structures might possibly indicate disorder in the solvent channels and information might be missing to analyse possible steric clashes. The importance of steric clashes has previously been investigated in multiple systems to justify binding preferences. For example, crystals of the Z-DNA forming sequences d(CGCGCG) and d(CGCGTG) which had been soaked in both Co (II) ions and Cu (II) ions were analysed. In B-DNA systems, steric clashes between the hydrated metals bound to the N7 position of guanines and the previous base were also reported to influence where the metals would bind³². Likewise, the observation that bound guanines are often located at the end of an oligonucleotide in duplexes containing Ni (II) ions would often be explained by the reduction of clashes between the metal ion and the local environment^{15,33}. Furthermore, another study showed sequence-specific and strand-specific binding of divalent cations with a preference for GGCC sequences that could also be explained by the hydration sphere around the metal ion as the cause of steric clashes³⁴. However, some of these differences could be explained by the different methods which can be used to introduce transition metals into the system, since all the structures used as examples here would have been acquired from crystals where the metal ions were introduced by either cocrystallisation or soaking. The differences between these two methods and the potential consequences on metal binding, metal availability for reductants and preferred metal binding sites were discussed in Chapter 2. Still, the presence of an ordered bulk solvent to be able to identify the first shell, or the molecules and atoms directly coordinated to the metal ion, and second shell of interactions is still important regardless of the method used^{35,36}.

Demonstrations of binding preferences such as reported in this study could suggest damage preferences in the event the bound base is the one which ends up oxidised by the hydroxyl radicals. In this case, the three structures might illustrate three potential intermediates in a Fenton-like copper environment. Initially, and as illustrated in **8BAE**, copper (II) ions are

found coordinated to the guanine bases at the 5' end of the strands and, supposing the closest base gets damaged during the reaction between copper and hydrogen peroxide, this guanine would be oxidised to 8-oxoguanine, an intermediate represented by **8BAF**. The crystals of the damaged sequence were grown in the same conditions as the native sequence and a similar non-canonical structure was solved. Furthermore, the change from G to oxoG did not influence how the sequence crystallised. However, metal ions need a donor atom to have one pair of electrons available to be able to directly coordinate and the nitrogen in the N7 position can no longer donate a pair of electrons after guanine oxidation (Figure 3.7). Ultimately, this would prevent copper (II) ions from coordinating to the guanine bases at the 5' end of the strands after oxidative damage at this position. Finally, **8BAG** would represent the last intermediate, when the damaged base, oxoG1, is present and copper (II) ions are still available to coordinate. Surprisingly, coordinated copper (II) ions have been located at the N7 position of G5, a guanine which previously had no copper (II) ions bound in **8BAE**.

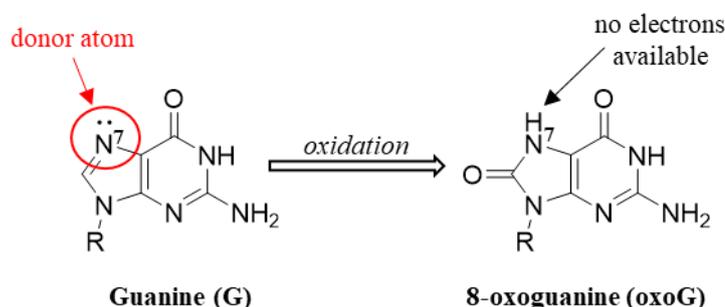


Figure 3.7: Oxidation reaction of guanine to form 8-oxoguanine. A nitrogen (N7, circled in red) can donate its pair of electrons to directly coordinate to metal ions. After oxidation, these electrons are no longer available and metal ions can no longer coordinate directly to N7.

The results were surprising, as the absence of a metal centre at the N7 position of G5 and G12 was expected in **8BAG**. No copper (II) ions were found to coordinate to the N7 position of G5 in **8BAF**, and since the two structures were found to be nearly identical after the RMSD was calculated, the same was predicted to happen in **8BAG**. However, this is not what has been reported, and two copper (II) ions were found bound to G5 and G12. Furthermore, no major differences were identified between the local environments around the two neighbouring guanines, which includes waters and hexamine cobalt complexes. However, disorder in the solvent channel in all structures could potentially explain why no potential explanations for

steric clashes have been identified. In DNA crystals, an ordered hydrogen bond network of water molecules would usually be found, a network more apparent the more ordered water molecules are placed at higher resolution³⁷, something that has not been observed in all the three structures here despite the high-resolution. Furthermore, low occupancy of the metal centres and a partial hydration sphere in both **8BAE** and **8BAG** could also explain why some information is missing to explain why the 5' position is favoured in the absence of the damaged base. Nonetheless, this does not explain why the second guanine is suddenly bound and has a metal centre following oxidative damage if this was an unfavourable in the native environment. Yet, other different hypothesis could be considered. As discussed in Chapter 2, a concentration in metal ions too low or short soaking time can influence binding preferences, especially if the concentration in copper (II) ions is too low relative to the concentration in DNA in the crystal. But this scenario was anticipated and the crystals for **8BAE** and **8BAG** were soaked for five days and one day respectively, until the crystals were too damaged to diffract or dissolved. Furthermore, in addition to steric clashes, charge effects might influence binding preferences in this case. Remarkably, DNA are negatively charged molecules which attract positively charged ions³⁸. Yet, it might be possible that the high solvent content in hexammine cobalt (III) ions and copper (II) ions, coupled with the disorder, has influenced binding preferences. A theory which could also explain why less hexammine cobalt were identified in the asymmetric unit of **8BAE** after soaking. However, although why G5 and G12 were not a preferred binding sites for copper is debatable, the results can suggest a potential mechanism for oxidative DNA damage in Fenton-like copper system. With the hypothesis oxidative damage might occur where the copper binds to the DNA, the possibility that not all guanines are occupied by a copper means the likelihood of specific positions being damaged is greater. However, once damaged, the damaged guanine can no longer coordinate with a copper (II) ion and a new preferred binding sites is occupied on a previously otherwise unoccupied guanine. Then, this guanine could be damaged and in turn, new possible binding sites might be preferred, generating even more oxidative damage. Damage progression in Fenton systems is not a mechanism that has been previously described. Yet, it is still important to consider to be able to understand how DNA damage can progress in oxidative stress environment, especially at an atomic level. Additionally, this might also indicate that information on preferred binding sites alone might not be sufficient to predict hot spots for oxidative damage in DNA. Nevertheless, soaking crystals containing copper (II) ions with hydrogen peroxide, and potentially observe

oxidative damage after an in-crystal reaction, would help to test whether or not damaged is contained to bases adjacent to the reaction centre.

3.5. References

1. Twomey, P. J., Viljoen, A., House, I. M., Reynolds, T. M. & Wierzbicki, A. S. Relationship between Serum Copper, Ceruloplasmin, and Non-Ceruloplasmin-Bound Copper in Routine Clinical Practice. *Clin. Chem.* **51**, 1558–1559 (2005).
2. Lin, T.-Y. & Wu, C.-H. Activation of hydrogen peroxide in copper(II)/amino acid/H₂O₂ systems: effects of pH and copper speciation. *J. Catal.* **232**, 117–126 (2005).
3. Perez-Benito, J. F. Reaction pathways in the decomposition of hydrogen peroxide catalyzed by copper(II). *J. Inorg. Biochem.* **98**, 430–438 (2004).
4. Moffett, J. W. & Zika, R. G. Reaction kinetics of hydrogen peroxide with copper and iron in seawater. *Environ. Sci. Technol.* **21**, 804–810 (1987).
5. Millero, F. J., Sharma, V. K. & Karn, B. The rate of reduction of copper(II) with hydrogen peroxide in seawater. *Mar. Chem.* **36**, 71–83 (1991).
6. Pham, A. N., Xing, G., Miller, C. J. & Waite, T. D. Fenton-like copper redox chemistry revisited: Hydrogen peroxide and superoxide mediation of copper-catalyzed oxidant production. *J. Catal.* **301**, 54–64 (2013).
7. Hofer, T. & Perry, G. Nucleic acid oxidative damage in Alzheimer's disease—explained by the hepcidin-ferroportin neuronal iron overload hypothesis? *J. Trace Elem. Med. Biol.* **38**, 1–9 (2016).
8. Sliwinska, A. *et al.* The levels of 7,8-dihydrodeoxyguanosine (8-oxoG) and 8-oxoguanine DNA glycosylase 1 (OGG1) – A potential diagnostic biomarkers of Alzheimer's disease. *J. Neurol. Sci.* **368**, 155–159 (2016).
9. Valavanidis, A., Vlachogianni, T. & Fiotakis, C. 8-hydroxy-2'-deoxyguanosine (8-OHdG): A Critical Biomarker of Oxidative Stress and Carcinogenesis. *J. Environ. Sci. Health Part C* **27**, 120–139 (2009).
10. Nakae, Y., Stoward, P. J., Bessalov, I. A., Melamede, R. J. & Wallace, S. S. A new technique for the quantitative assessment of 8-oxoguanine in nuclear DNA as a marker of oxidative stress. Application to dystrophin-deficient DMD skeletal muscles. *Histochem. Cell Biol.* **124**, 335–345 (2005).

11. Collins, A. Comparison of different methods of measuring 8-oxoguanine as a marker of oxidative DNA damage. *Free Radic. Res.* **32**, 333–341 (2000).
12. Fleming, A. M., Muller, J. G., Ji, I. & Burrows, C. J. Characterization of 2'-deoxyguanosine oxidation products observed in the Fenton-like system Cu(II)/H₂O₂/reductant in nucleoside and oligodeoxynucleotide contexts. *Org. Biomol. Chem.* **9**, 3338–3348 (2011).
13. Saito, I. *et al.* Photoinduced DNA Cleavage via Electron Transfer: Demonstration That Guanine Residues Located 5' to Guanine Are the Most Electron-Donating Sites. *J. Am. Chem. Soc.* **117**, 6406–6407 (1995).
14. Kagawa, T. F., Geierstanger, B. H., Wang, A. H. & Ho, P. S. Covalent modification of guanine bases in double-stranded DNA. The 1.2-Å Z-DNA structure of d(CGCGCG) in the presence of CuCl₂. *J. Biol. Chem.* **266**, 20175–20184 (1991).
15. Rohner, M., Medina-Molner, A. & Spingler, B. N,N,O and N,O,N Meridional cis Coordination of Two Guanines to Copper(II) by d(CGCGCG)₂. *Inorg. Chem.* **55**, 6130–6140 (2016).
16. Atwell, S., Meggers, E., Spraggon, G. & Schultz, P. G. Structure of a Copper-Mediated Base Pair in DNA. *J. Am. Chem. Soc.* **123**, 12364–12367 (2001).
17. Govindaraju, M. *et al.* Copper interactions with DNA of chromatin and its role in neurodegenerative disorders. *J. Pharm. Anal.* **3**, 354–359 (2013).
18. Fritzsche, H. & Zimmer, C. Infrared Studies of Deoxyribonucleic Acids, their Constituents and Analogues. *Eur. J. Biochem.* **5**, 42–44 (1968).
19. McCoy, A. J. *et al.* Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674 (2007).
20. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2126–2132 (2004).
21. Afonine, P. V. *et al.* Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr. D Biol. Crystallogr.* **68**, 352–367 (2012).

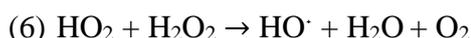
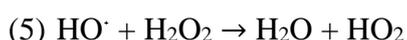
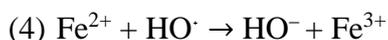
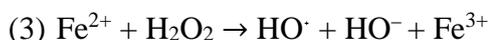
22. Thorpe, J. H., Teixeira, S. C. M., Gale, B. C. & Cardin, C. J. Crystal structure of the complementary quadruplex formed by d(GCATGCT) at atomic resolution. *Nucleic Acids Res.* **31**, 844–849 (2003).
23. Moriarty, N. W., Grosse-Kunstleve, R. W. & Adams, P. D. electronic Ligand Builder and Optimization Workbench (eLBOW): a tool for ligand coordinate and restraint generation. *Acta Crystallogr. D Biol. Crystallogr.* **65**, 1074–1080 (2009).
24. Winter, G. xia2: an expert system for macromolecular crystallography data reduction. *J. Appl. Crystallogr.* **43**, 186–190 (2010).
25. Kabsch, W. XDS. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 125–132 (2010).
26. Schrödinger, LLC. The PyMOL Molecular Graphics System, Version 1.8. (2015).
27. Zheng, G., Lu, X.-J. & Olson, W. K. Web 3DNA—a web server for the analysis, reconstruction, and visualization of three-dimensional nucleic-acid structures. *Nucleic Acids Res.* **37**, W240–W246 (2009).
28. Macrae, C. F. *et al.* Mercury 4.0: from visualization to analysis, design and prediction. *J. Appl. Crystallogr.* **53**, 226–235 (2020).
29. Lee, H. *et al.* Activation of Oxygen and Hydrogen Peroxide by Copper(II) Coupled with Hydroxylamine for Oxidation of Organic Contaminants. *Environ. Sci. Technol.* **50**, 8231–8238 (2016).
30. Halliwell, B., Adhikary, A., Dingfelder, M. & Dizdaroglu, M. Hydroxyl radical is a significant player in oxidative DNA damage *in vivo*. *Chem. Soc. Rev.* **50**, 8355–8360 (2021).
31. Barone, G. *et al.* DNA-binding of nickel(II), copper(II) and zinc(II) complexes: Structure–affinity relationships. *Coord. Chem. Rev.* **257**, 2848–2862 (2013).
32. Gao, Y.-G., Sriram, M. & Wang, A. H.-J. Crystallographic studies of metal ion - DNA interactions: different binding modes of cobalt(II), copper(II) and barium(II) to N7 of guanines in Z-DNA and a drug-DNA complex. *Nucleic Acids Res.* **21**, 4093–4101 (1993).
33. Valls, N., Usón, I., Gouyette, C. & Subirana, J. A. A Cubic Arrangement of DNA Double Helices Based on Nickel–Guanine Interactions. *J. Am. Chem. Soc.* **126**, 7812–7816 (2004).

34. Chiu, T. K. & Dickerson, R. E. 1 Å crystal structures of B-DNA reveal sequence-specific binding and groove-specific bending of DNA by magnesium and calcium¹¹ Edited by I. Tinoco. *J. Mol. Biol.* **301**, 915–945 (2000).
35. Dudev, T., Lin, Dudev, M. & Lim, C. First–Second Shell Interactions in Metal Binding Sites in Proteins: A PDB Survey and DFT/CDM Calculations. *J. Am. Chem. Soc.* **125**, 3168–3180 (2003).
36. Levy, R., Sobolev, V. & Edelman, M. First- and second-shell metal binding residues in human proteins are disproportionately associated with disease-related SNPs. *Hum. Mutat.* **32**, 1309–1318 (2011).
37. Weichenberger, C. X., Afonine, P. V., Kantardjieff, K. & Rupp, B. The solvent component of macromolecular crystals. *Acta Crystallogr. D Biol. Crystallogr.* **71**, 1023–1038 (2015).
38. Lipfert, J., Doniach, S., Das, R. & Herschlag, D. Understanding Nucleic Acid–Ion Interactions. *Annu. Rev. Biochem.* **83**, 813–841 (2014).

Chapter 4 - Iron (II) Binding Preferences to DNA Structure – Observing Metal Coordination Before Oxidative Damage

4.1. Introduction

Iron (Fe) is an abundant transition metal involved in essential biological processes such as enzymatic functions¹, oxygen transport² and electron transport³. However, free iron (Fe (II)) can also have undesirable effects in the cellular environment by reacting with hydrogen peroxide (H₂O₂), a by-product of cellular respiration. This reaction, known as the Fenton reaction (Eq. (1)), generates a highly reactive oxygen species (ROS), hydroxyl radicals (HO•) which can damage cellular components⁴. The hydrogen peroxide decomposition mechanism was first suggested by Haber and Weiss in 1934 (Scheme 4.(1-4)) and has been used since then to describe hydroxyl radicals' production⁵.



Scheme 4.1: The reaction of iron with hydrogen peroxide

During normal cellular activity, antioxidants keep ROS levels balanced. However, an imbalance between ROS production and the ability to neutralise ROS in a cell, a phenomenon known as oxidative stress⁶, has been linked to many disorders, including neurological disorders⁷, cancer, and cardiovascular diseases⁸.

Nucleic acids have been shown to be easily damaged by ROS during oxidative stress generated by elevated Fe²⁺ concentrations. For example, an increased level of oxidative damage reflected by the presence of the oxidative stress marker product 8-oxo-7,8-dihydro-2'-

deoxyguanine (oxoG), has been reported in the brains of individuals having suffered from neurodegenerative diseases, such as Alzheimer's and Parkinson's diseases, and has been linked to iron dysregulation⁹. Moreover, the ability of iron to bind to DNA through the N7 position of purine bases and the phosphate backbone, and the short-lived existence of hydroxyl radicals, suggests that damage can be correlated to binding preferences of iron¹⁰⁻¹². And often, this correlation has been used to discuss sequence-specific DNA damages including oxoG and double-stranded cleavage. Recently, Fleming *et al.* has conducted multiple studies to identify oxidation products after using copper (II) and iron (II) to catalyse the Fenton reaction^{13,14}. It has been concluded that copper coordination to guanine played a major role in base oxidation, while iron coordination to the phosphate backbone could explain the prevalence of strand-break in Fe (II)/H₂O₂ systems. However, few studies have looked in details on binding preferences of Fe (II) in the presence of DNA as a function of structure. 20 years ago, a study by Rai *et al.* used NMR to confirm iron binding to TG steps with neighbouring purines at both 5' and 3' extremities (5'-Pu-TGR-Pu-3'), and the result determined that Fe (II) interacts preferentially at the 5'-ATGA-3' site through the N7 position of the guanine¹⁵. This system, identified by Henle *et al.*, has shed some light on iron positioning and helped understand why specific strand-cleavage observed upstream to this guanine might occur¹⁰. However, since then, no structures of DNA in the presence of iron have been reported to further investigate iron-binding preferences. Furthermore, while the strong correlation between iron binding sites and the type of DNA damage that occurs has been brought up as an explanation for preferential DNA damage, at the present, no experiment has been conducted to look at damage before and immediately after identifying a damage hot-spot.

Binding preferences of many transition metals to DNA, such as to copper, cobalt, manganese, and zinc, have been identified using X-Ray crystallography, as observed in the previous Chapters of this thesis. These studies have been used as the basis to understand binding preferences of iron, as systems containing iron (II) can be challenging. Iron (II) oxidises easily in an aerobic environment, and depending on pH, only a couple of hours are necessary to convert all iron (II) ions to iron (III) ions¹⁶. Crystal growth has been known to sometimes take several months, and even soaking experiments can take several days, thus an anoxic environment would be necessary to obtain DNA crystals containing ordered iron (II) ions.

To identify damage hot-spots, the chosen crystal system not only would need to withstand metal soaking, but it would also need to withstand hydrogen peroxide. In this study, a crystal system has been selected to attempt iron (II) chloride soaking and obtain a DNA structure with ordered metals. The Z-DNA forming sequence d(CGCGCG) and its structure have been extensively studied using X-Ray crystallography, notably by Brzezinski *et al.*, whose work reported the differences in crystal packing found in d(CGCGCG) structures, described as the polymorphic form A and polymorphic form B¹⁷. Additionally, this sequence has been shown to form robust crystals that can diffract up to 0.6 Å resolution, which have been used previously to obtain accurate characteristics of metal coordination after soaking. These characteristics highlighted this crystal system as a perfect candidate to withstand both metal soaking and reactant soaking. In this chapter, the results have demonstrated six potential reaction centres in the presence of six DNA-bound copper (II) ions. However, the results also highlight the potential differences between copper and iron interactions with DNA, as only four iron (II) ions were found to be bound to DNA. These differences could help understand the disparity observed between DNA oxidative damage caused by copper-catalysed Fenton and iron-catalysed Fenton chemistry.

4.2. Material and methods

4.2.1. DNA preparation

The oligonucleotide d(CGCGCG) was purchased from Eurogentec as a solid purified by RP-HPLC. DNA concentration was determined from the value of the absorbance at $\lambda = 260$ nm, using the molar extinction coefficient provided by the manufacturer. 1.5 mM solution of the DNA was annealed at 358 K for two minutes then allowed to cool slowly to room temperature.

4.2.2. Crystallisation and soaking experiment

Single crystals to be used for soaking were grown at 291 K by the sitting-drop vapor diffusion method using the protocol¹⁸ described by Drozdal *et al.* The drops contained 2 μL of 1.5 mM oligonucleotides and 2 μL of precipitating solution containing 10 % (v/v) 2-methyl-2,4-pentanediol (MPD), 40 mM sodium cacodylate, pH 6.0, 80 mM KCl, 12 mM NaCl and 12 mM spermine tetrachloride. The drops were equilibrated against 0.5 ml of 35 % (v/v) MPD. The crystals appeared within 48 hours.

To ensure the iron (II) chloride present in the solutions does not oxidise in the presence of oxygen during the soaking experiment, a plate containing crystals grown in an aerobic environment, as well as pre-weighted iron (II) chloride, water, and a pre-mixed crystallisation solution, were transferred inside a glove bag with a constant supply of nitrogen (N_2) and left for 24 hours to equilibrate. Prior to soaking, the pre-weighted iron (II) chloride was mixed with water to a concentration of 10 mM. The plate was kept in the glove bag while single crystals were transferred from their drop to a drop containing 2 μL of water, 2 μL of precipitating solution and 4 μL of 10 mM iron (II) chloride, to a total of 8 μL . The crystals were left to soak in anoxic conditions for 48 hours. During iron-soaking, the crystals' colour appeared to change to a light blue colour (Figure 4.1). This change was used to track the experiment.

Crystals soaked in copper (II) chloride did not require any adjustments. The crystals were transferred to drops containing 2 μL of precipitating solution and 2 μL of copper 10 mM (II) chloride for 72 hours. Similarly, the crystal's colour change to light blue was monitored to track the experiment.

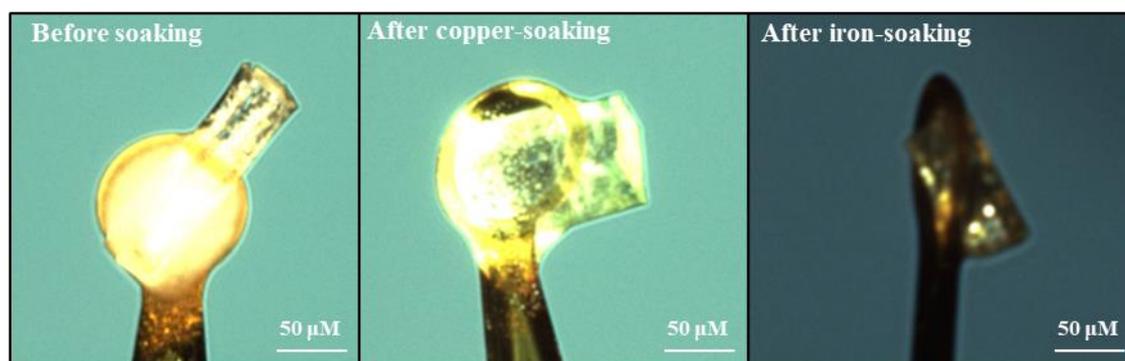


Figure 4.1: Crystals of the DNA sequence d(CGCGCG) before (left) and after the crystals were soaked in drops containing 10 mM CuCl_2 (middle) and 10 mM FeCl_2 . The native crystals were uncoloured. After soaking in both copper (II) and iron (II), the crystal would take a light blue colour. These changes were used to monitor the experiment.

4.2.3. Data collection and processing

Data from the native, copper-soaked and iron-soaked crystals were collected from single crystals of approximate dimensions 200 x 100 x 100 μm , at 100K using a XtaLAB Synergy-S diffractometer and a PhotonJet-S X-ray source at a wavelength of 1.5406 \AA with a HyPix-6000HE detector. The data collection and refinement statistics of all three datasets can be found in Table 4.1.

2760 frames were collected from the native crystal. The data were integrated and scaled using the CrysAlisPRO software, giving 10952 unique reflections. The structure was solved in $\text{P2}_1\text{2}_1\text{2}_1$ by molecular replacement using Phaser¹⁹ and the structure **1I0T** from the Protein Data Bank. The model was updated using Coot²⁰ and refined using phenix.refine²¹. To confirm that the model was satisfactory and that the observed data and the calculated data from the model were in agreement, the R_{free} and the R_{work} were calculated after each round of refinement. R-factors can be made arbitrarily low by overfitting the data during refinement. Consequently, five percent of the reflections were reserved for the R_{free} set and not used during refinement. The R_{free} and R_{work} observed after the final round of refinement for this model were 0.1252 of 0.1053 respectively. These values can be compared relatively to X-ray structures of similar resolution after deposition and validation in the Protein Data Bank. The data, and final coordinates, were deposited in the Protein Data Bank with ID **8OEX**.

1350 frames were collected from the copper-soaked crystal. The data were also integrated and scaled using the CrysAlisPRO software, giving 7930 unique reflections. The structure was solved in $P2_12_12_1$ by molecular replacement using Phaser¹⁹ and the structure **1I0T** from the Protein Data Bank. The model was updated using Coot²⁰ and refined using phenix.refine²¹ to give a final R_{free} of 0.1458 and R_{work} of 0.1100. Five percent of reflections were reserved for the R_{free} set. The data, and final coordinates, were deposited in the Protein Data Bank with ID **8OEY**.

Finally, to transfer the iron-soaked crystal from the glove bag to the goniometer for data collection, the crystal had to be collected with a loop from the plate still present in the glove bag, then transferred from the bag to the goniometer. 1314 frames were collected from the iron-soaked crystal. The data were also integrated and scaled using the CrysAlisPRO software, giving 3119 unique reflections. The structure was solved in $P2_12_12_1$ by molecular replacement using Phaser¹⁹ and the structure **1I0T** from the Protein Data Bank. The model was updated using Coot²⁰ and refined using phenix.refine²¹ to give a final R_{free} of 0.2263 and R_{work} of 0.1947. Five percent of reflections were reserved for the R_{free} set. The data, and final coordinates, were deposited in the Protein Data Bank with ID **8OEZ**.

Table 4.1: Data collection and refinement statistics for 8OEX, 8OEY and 8OEZ. Outer shell statistics shown in parentheses.

PDB ID	8OEX	8OEY	8OEZ
Data collection			
Space group	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁
Cell dimensions			
<i>a</i> , <i>b</i> , <i>c</i> , (Å)	18.28 30.76 42.48	18.17 30.46 43.00	18.09 30.39 42.51
Resolution, (Å)	18.28 - 1.07 (1.09 - 1.07)	18.17 - 1.20 (1.22 - 1.20)	14.60 - 1.64 (1.70 - 1.64)
R _{merge}	0.189 (0.576)	0.058 (0.346)	0.105 (0.197)
R _{meas}	0.198 (0.606)	0.062 (0.377)	0.123 (0.223)
R _{pim}	0.057 (0.180)	0.020 (0.148)	0.039 (0.073)
Total number of observations	133444	72414	30249
Total number of unique observations	10952	7930	3119
CC _{1/2}	0.999 (0.868)	1.000 (0.932)	1.000 (0.996)
I/σI	11.4 (3.4)	22.9 (5.1)	13.7 (8.0)
Completeness, (%)	99.6 (92.7)	99.9 (99.4)	99.3 (93.8)
Multiplicity	12.2 (9.7)	9.1 (6.3)	9.7 (8.8)
Refinement			
Resolution, (Å)	17.478 - 1.075 (1.113 - 1.075)	17.566 - 1.200 (1.243 - 1.200)	14.600 - 1.642 (1.701 - 1.642)
No. Reflections	10894 (1019)	7883 (747)	3084 (271)
R _{free} /R _{work}	0.1053/0.1252	0.1100/0.1458	0.2263/0.1947
rmsd			
Bond lengths, (Å)	0.011	0.014	0.017
Bond angles, (°)	1.40	1.65	1.69

4.2.4. Structural and packing analysis.

All figures were created using PyMOL²² unless specified otherwise. Dinucleotide step parameters were calculated using Web 3DNA 2.0²³. CCP4 maps were generated from the final mtz file, including the anomalous map when available, using the phenix.mtz2map program. Maps were visualised in PyMOL and displayed using the mesh function. *2mFo-DFc* maps were contoured at 1σ unless specified otherwise. Anomalous maps were contoured at 2σ unless specified otherwise. Packing diagrams were generated using Mercury²⁴.

4.3. Results

4.3.1. DNA hexamer: overall structure and helical parameters analysis

8OEX was obtained from a native crystal of the DNA hexamer sequence d(CGCGCG) (Figure 4.2). The crystal diffracted to a high resolution of 1.07 Å after collection on an in-house X-Ray source. Although this structure will not show the level of details and accuracy other structures from the same system have demonstrated previously¹⁷, this structure can be directly compared to the structures displaying copper and iron binding preferences after soaking.

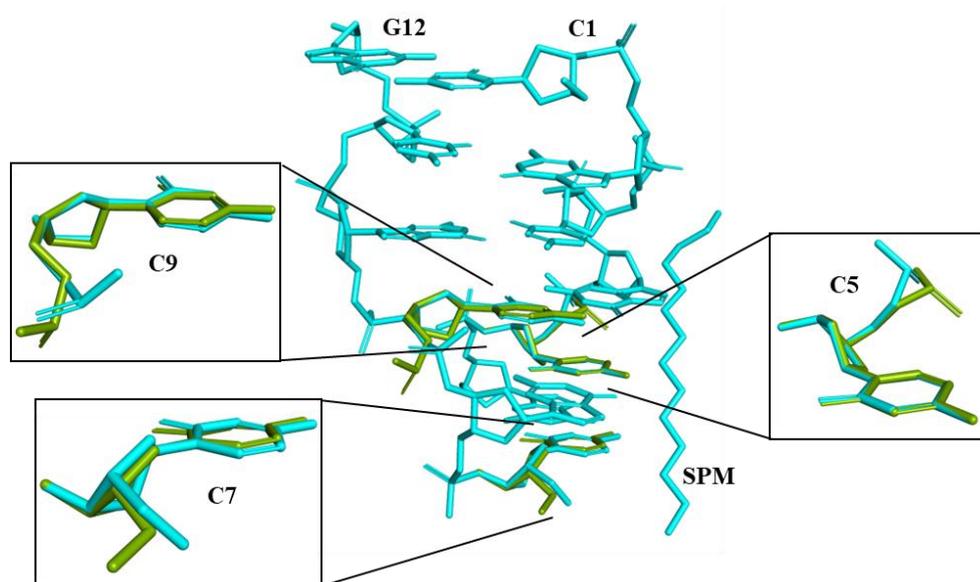


Figure 4.2: Native structure of the DNA hexamer sequence d(CGCGCG). The alternative conformations are shown in cyan (A) and green (B). The duplex is shown as sticks.

The asymmetric unit of **8OEX** contains two strands for a total of 12 nucleotides and forms a duplex. A spermine molecule was also found to interact with the two Z-DNA duplexes in the crystal lattice, in a similar way to the other Z-DNA structures presenting the same crystal packing. The nucleotides from chain A are labelled from C1 to G6 in the 5'-3' direction and the nucleotides from chain B are labelled from C7 to G12 in the 5'-3' direction. The structural parameters classify the duplex as a left-handed Z-form structure. Disorder is observed at the G4-C5 and G8-G9 phosphate linkages, as well as at the C7 5'-OH group. The alternative conformations are referred to as A and B, and have occupancies of 69/31, 57/43 and 46/54 % respectively (Figure 4.2). Details of the local base-pair parameters can be found in the Supplementary Information. The pseudorotation angles between the alternative conformations

were found to slightly vary, but not enough to change sugar pucker. The structure belongs to the space group $P2_12_12_1$, and crystal packing is consistent with the polymorphic form B of d(CGCGCG) Z-DNA structures (Figure 4.3).

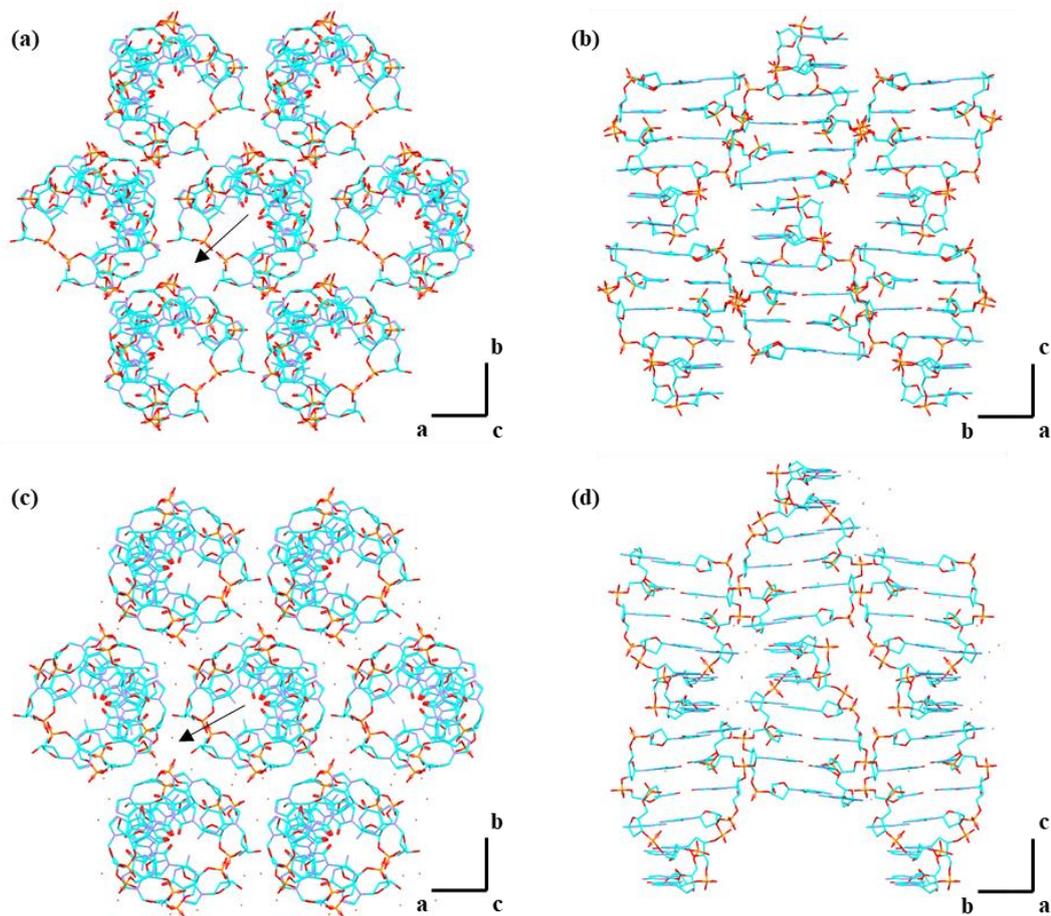


Figure 4.3: Packing diagram of the 8OEX structure (a and b) compared to the packing diagram of the 1D39 structure (c and d) with views along the c-axes (a and c) and the a-axis (b and d). This highlights the differences between the two different mode of crystal packing observed in d(CGCGCG) systems, also described as polymorphic form A and polymorphic form B.

8OEY, the copper-soaked structure was obtained from a native crystal of the DNA sequence d(CGCGCG) grown in the same conditions as the unsoaked structure. After soaking the crystal in copper (II) chloride for 72 hours, the crystal diffracted to 1.20 Å. The asymmetric unit contains two strands for a total of 12 nucleotides and forms a duplex as well. Nucleotides are labelled in the same way as the native structure. The nucleotides from chain A are labelled from C_{Cu1} to G_{Cu6} in the 5'-3' direction and the nucleotides from chain B are labelled from

C_{Cu}7 to G_{Cu}12 in the 5'-3' direction. Soaking did not change the overall structure parameters and the duplex is still classified as a left-handed Z-form DNA. A comparison between this structure and the native structure was conducted and a root-mean-square deviation (RMSD) of 0.492 Å was calculated, which indicates high consistency (Figure 4.4a). Disorder was also found at the C_{Cu}7 5'-OH group, as well as at the G_{Cu}4-C_{Cu}5 phosphate linkage. The alternative conformations are referred to as C and D and have occupancies of 58/42 and 52/48 % respectively. Details of the local base-pair parameters can be found in the Supplementary Information. Additionally, different sugar puckers were found at C_{Cu}5, with conformation C and D in the C1'-exo and C2'-endo range. In comparison, the same nucleotide was in the C1'-exo range in the native structure. The sugar puckers at G_{Cu}4 and G_{Cu}8 also differs from the equivalent nucleotides in the native structure, with pseudorotation angles of 36.0° and 24.9°, classifying both as C3'-endo and C4'-exo respectively. Comparison of sugar puckering can be found in Table 4.2.

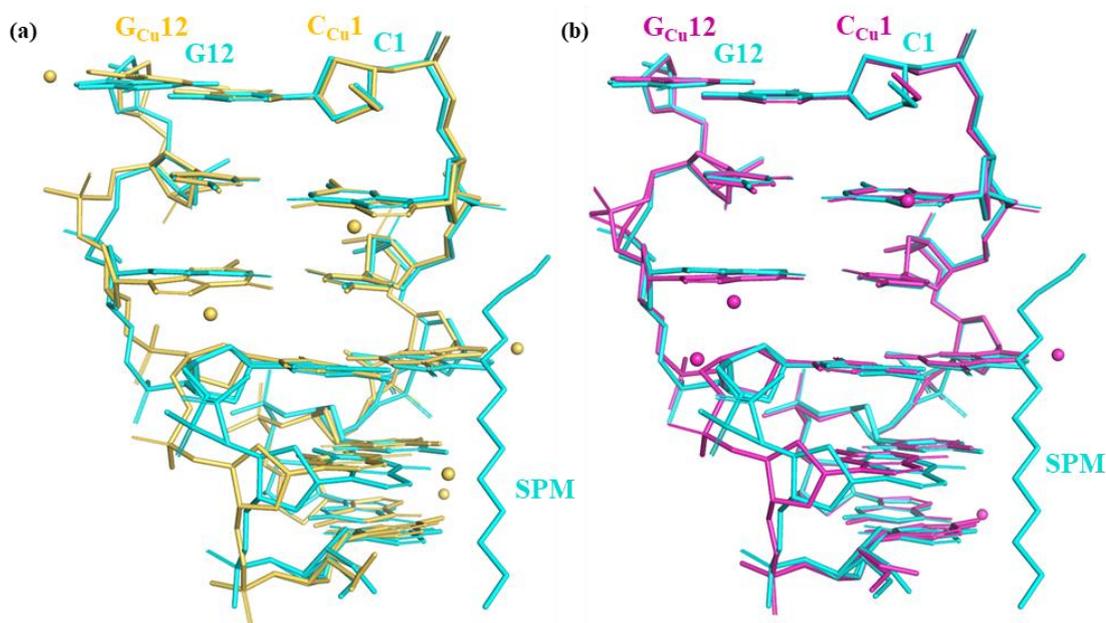


Figure 4.4: a. Superimposition of 8OEX (cyan) and 8OEY (orange). b. Superimposition of 8OEX (cyan) and 8OEZ (purple). Metal ions are represented as spheres.

8OEZ, the iron-soaked structure, was obtained from a native crystal of the DNA sequence d(CGCGCG) grown in the same conditions as the unsoaked structure but kept in a glove bag containing N₂ before attempting soaking, to prevent oxygen to be present in the system. The crystal diffracted to 1.64 Å. Like with **8OEX** and **8OEY**, the asymmetric unit contains two strands for a total of 12 nucleotides and forms a duplex. The same labelling system is also used here as well. The nucleotides from chain A are labelled from C_{Fe}1 to G_{Fe}6 in the 5'-3' direction and the nucleotides from chain B are labelled from C_{Fe}7 to G_{Fe}12 in the 5'-3' direction. Soaking did not change the overall structure parameters and the duplex is still classified as a left-handed Z-form DNA. A comparison between this structure and the native structure was also drawn and a RMSD of 0.254 Å was calculated, which indicates high consistency between these two structures (Figure 4.4b). Disorder was similar and found at the C_{Fe}7 5'-OH group, as well as the G_{Fe}10-C_{Fe}11 phosphate linkage. The alternative conformations are referred to as E and F and have occupancies of 62/38 and 53/47 % respectively. Details of the local base-pair parameters can be found in the Supplementary Information. Like the copper-soaked structure, different sugar puckers were found at C_{Fe}7, with conformation E and F in the C1'-exo and C2'-endo range. However, unlike **8OEY**, many sugar puckers in this structure differs from the native structure. The sugar puckers at G_{Fe}2, G_{Fe}4, C_{Fe}5 and G_{Fe}10 were classified as C4'-exo, C4'-exo, C2'-endo and C4'-exo respectively. Comparison of sugar puckering can be found in Table 4.2. It has also been noted that no electron density associated with the spermine molecule was observed after both soaking experiments.

Table 4.2: Sugar pucker pseudorotation angles and sugar pucker type for the oligonucleotide d(CGCGCG) in 8OEX, 8OEY and 8OEZ. Upper/bottom values refer to two alternative conformations present. Which conformation the value refers to is shown in parentheses.

	Base	8OEX		8OEY		8OEZ	
		Phase angle of pseudorotation (P), °	Sugar Pucker	Phase angle of pseudorotation (P), °	Sugar Pucker	Phase angle of pseudorotation (P), °	Sugar Pucker
C H A I N A	5'-C	147.8	C2'-endo	155.8	C2'-endo	153.2	C2'-endo
	G	35.2	C3'-endo	26.5	C3'-endo	39.6	C4'-exo
	C	165.0	C2'-endo	166.7	C2'-endo	161.9	C2'-endo
	G	35.7	C3'-endo	36.0	C4'-exo	40.9	C4'-exo
	C	141.1 (A)	C1'-exo	143.4 (C)	C1'-exo	145.9	C2'-endo
			141.0 (B)	C1'-exo	151.8 (D)	C2'-endo	
	G-3'	168.3	C2'-endo	165.6	C2'-endo	163.3	C2'-endo
C H A I N B	5'-C	162.9 (A)	C2'-endo	171.5 (C)	C2'-endo	142.7 (E)	C1'-exo
		162.4 (B)	C2'-endo	159.5 (D)	C2'-endo	166.3 (F)	C2'-endo
	G	36.5	C4'-exo	24.9	C3'-endo	45.4	C4'-exo
	C	153.9 (A)	C2'-endo	158.1	C2'-endo	144.4	C2'-endo
		162.1 (B)	C2'-endo				
	G	25.7	C3'-endo	31.7	C3'-endo	45.5	C4'-exo
	C	156.3	C2'-endo	150.1	C2'-endo	146.4 (E)	C2'-endo
					148.9 (F)	C2'-endo	
	G-3'	168.3	C2'-endo	166.1	C2'-endo	166.5	C2'-endo

4.3.2. Cu²⁺ and Fe²⁺ ions coordination analysis

A total of six copper (II) ions were found in **8OEY** after the native crystal was soaked in copper (II) chloride. The Cu²⁺ ions were added based on the electron density present and validated based on the presence of peaks in the anomalous map at 2σ (Figure 4.5a). All Cu²⁺ ions were coordinated to the N7 position of guanines. After occupancy refinement, Cu²⁺ (1), Cu²⁺ (2), Cu²⁺ (3), Cu²⁺ (4), and Cu²⁺ (5) have partial occupancies of 51, 57, 51, 48, and 52 % respectively. Cu²⁺ (6) was modelled as having alternative conformations (A and B) due to the overall geometry and its proximity to the guanine and Cu²⁺ (5) (Figure 4.6a). Cu²⁺ (6A) and Cu²⁺ (6B) have partial occupancies of 32 and 17% respectively. The coordination sphere of Cu²⁺ (1), Cu²⁺ (2), Cu²⁺ (3), Cu²⁺ (4) and Cu²⁺ (5) all contain five water molecules (Figure 4.6a), while it is difficult to attribute water molecules to the coordination spheres of Cu²⁺ (6A) and Cu²⁺ (6B). However, after considering partial occupancies and distances, three water molecules were attributed to the coordination spheres of Cu²⁺ (6A) and four water molecules to the coordination spheres of Cu²⁺ (6B) (Figure 4.6a). A close-up of all the guanines present in the

asymmetric unit can be seen in Figure 4.8. Details of copper coordination are available in Table 4.3.

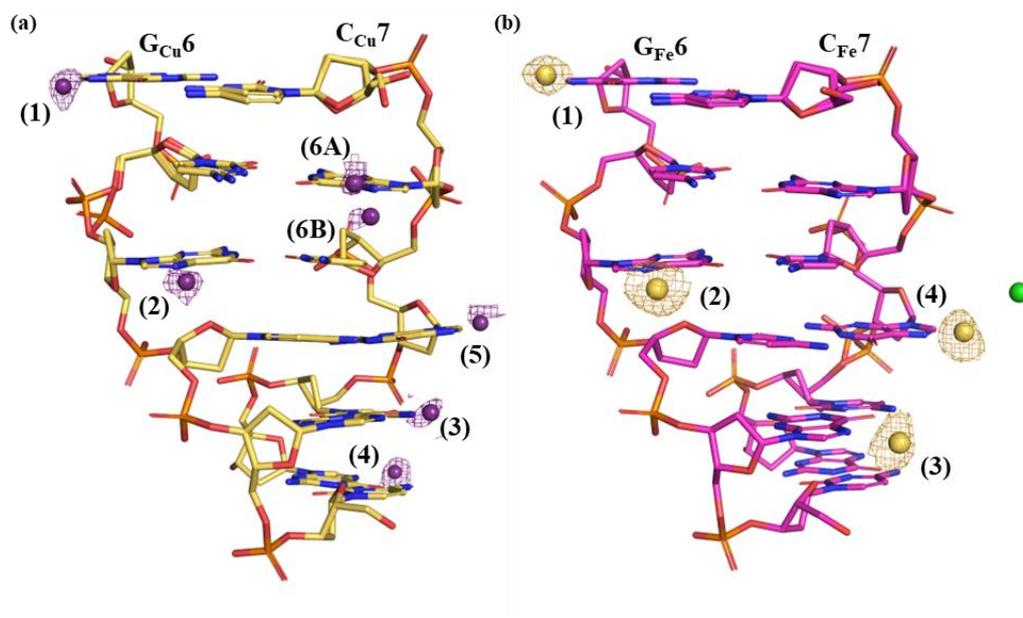


Figure 4.5: a. Structure of 8OEY with anomalous difference map for the six Cu^{2+} cations. Carbon, oxygen, nitrogen, and phosphate atoms are shown as yellow, red, blue, and orange respectively. Copper ions are shown as purple spheres. b. Structure of 8OEZ with anomalous difference map for the four Fe^{2+} cations. Carbon, oxygen, nitrogen, and phosphate atoms are shown as magenta, red, blue, and orange respectively. Iron ions are shown as yellow spheres. Chloride ions are shown as green spheres.

A total of four iron (II) ions were found in **8OEZ** after a native crystal was soaked in iron (II) chloride. The Fe^{2+} ions were added based on the electron density present and validated based on the presence of peaks in the anomalous map at 2σ (Figure 4.5b). All Fe^{2+} ions were coordinated to the N7 position of guanines. After occupancy refinement, Fe^{2+} (2), Fe^{2+} (3) and Fe^{2+} (4) remained at a 100 % occupancy. Fe^{2+} (1) has a partial occupancy of 74 %. The coordination sphere of Fe^{2+} (2) and Fe^{2+} (3) contain five water molecules (Figure 4.7b). However, both Fe^{2+} (1) and Fe^{2+} (4) have partial coordination spheres. Additionally, a chloride ion was modelled as part of the Fe^{2+} (4) ion's coordination sphere (Figure 4.6b). This was decided due to how far from the N7 position the density is and refinement did not give satisfactory results with other atoms modelled in this position. A close-up of all the guanines present in the asymmetric unit can be seen in Figure 4.8. Details of iron coordination are available in Table 4.3.

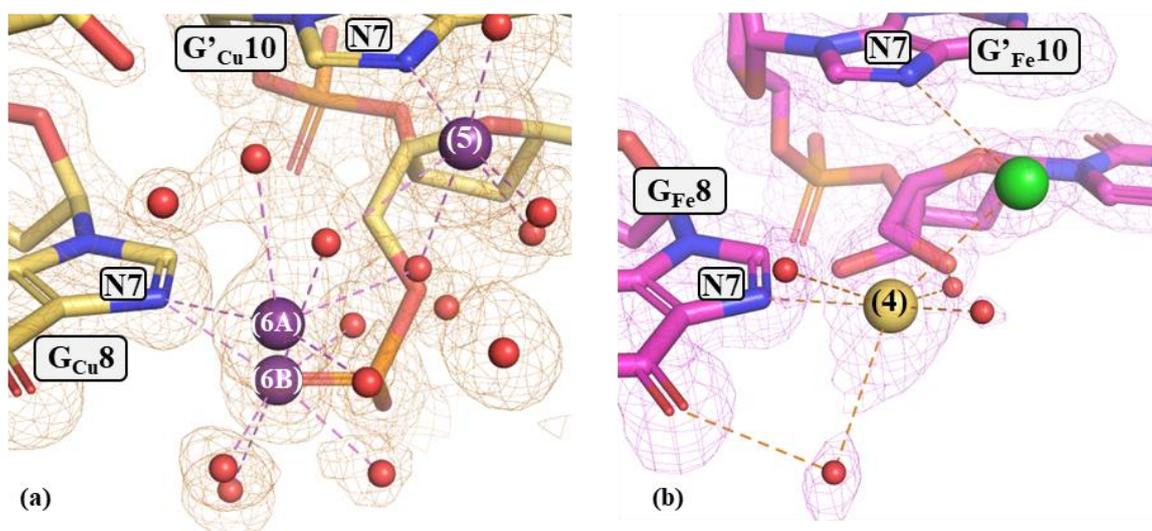


Figure 4.6: a. The coordination spheres of three Cu²⁺ (5, 6A and 6B) in 8OEY. Cu²⁺ (5) is shown here as coordinated to the guanine from a neighbouring duplex (G'_{Cu10}). The 2mFo-DFc electron density map is contoured at the 1.0σ level (orange). Carbon, oxygen, nitrogen, and phosphate atoms are shown as yellow, red, blue, and orange respectively. b. The coordination spheres of Fe²⁺ (4) in 8OEZ. Fe²⁺ (4) is shown here as coordinated to the guanine from a neighbouring duplex (G'_{Fe10}). The 2mFo-DFc electron density map is contoured at the 1.0σ level (purple). Carbon, oxygen, nitrogen, and phosphate atoms are shown as magenta, red, blue, and orange respectively. Iron ions are shown as yellow spheres. Chloride ions are shown as green spheres. Water molecules are shown as red spheres.

Table 4.3: Metals ions parameters from 8OEY and 8OEZ (distance in Å)

PDB ID	Metal	Occupancy	Distance to guanine N7	Distance to water molecules
8OEY	Cu ²⁺ (1)	0.51	1.89	1.93; 2.19; 1.92; 1.89; 2.38
	Cu ²⁺ (2)	0.57	2.01	2.28; 2.04; 2.00; 2.40; 1.94
	Cu ²⁺ (3)	0.48	1.98	2.12; 2.17; 1.86; 1.92; 2.16
	Cu ²⁺ (4)	0.52	2.00	1.87; 2.41; 1.97; 2.46; 1.93
	Cu ²⁺ (5)	0.51	1.97	2.27; 2.09; 1.86; 1.78; 2.29
	Cu ²⁺ (6A)	0.32	2.37	1.95; 2.45; 2.30
	Cu ²⁺ (6B)	0.17	2.05	1.93; 1.85; 1.83; 2.53
8OEZ	Fe ²⁺ (1)	0.74	2.35	2.24; 1.86
	Fe ²⁺ (2)	1.00	2.26	2.46; 2.23; 2.26; 2.31; 2.38
	Fe ²⁺ (3)	1.00	2.35	2.14; 2.07; 2.33; 1.99; 2.13
	Fe ²⁺ (4)	1.00	2.51	2.41; 1.89; 2.92; 3.04

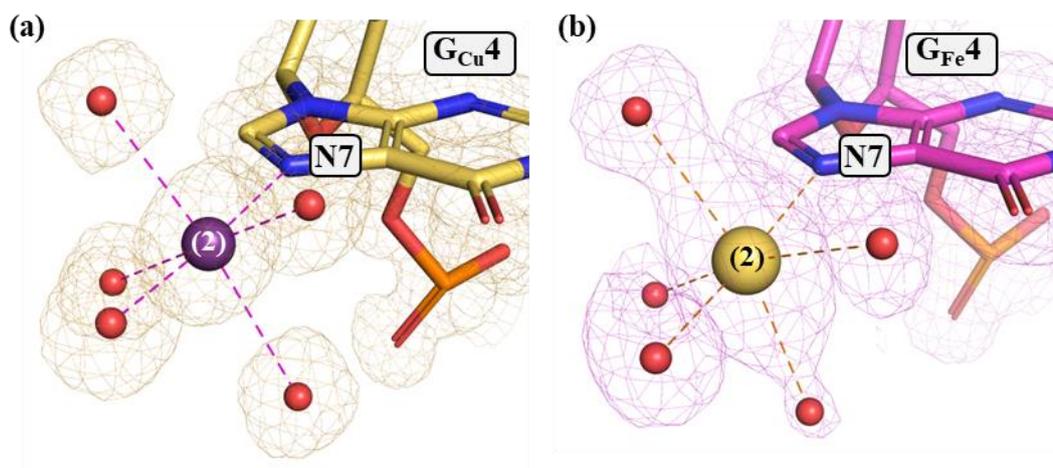


Figure 4.7: a. The coordination spheres of Cu^{2+} (2) in 8OEY. The $2mFo-DFc$ electron density map is contoured at the 1.0σ level (orange). Carbon, oxygen, nitrogen, and phosphate atoms are shown as yellow, red, blue, and orange respectively. Copper ions are shown as purple spheres. b. The coordination spheres of Fe^{2+} (2) in 8OEZ. The $2mFo-DFc$ electron density map is contoured at the 1.0σ level (purple). Carbon, oxygen, nitrogen, and phosphate atoms are shown as magenta, red, blue, and orange respectively. Iron ions are shown as yellow spheres. Both ions are coordinated to an equivalent guanine in the structure. Water molecules are shown as red spheres.

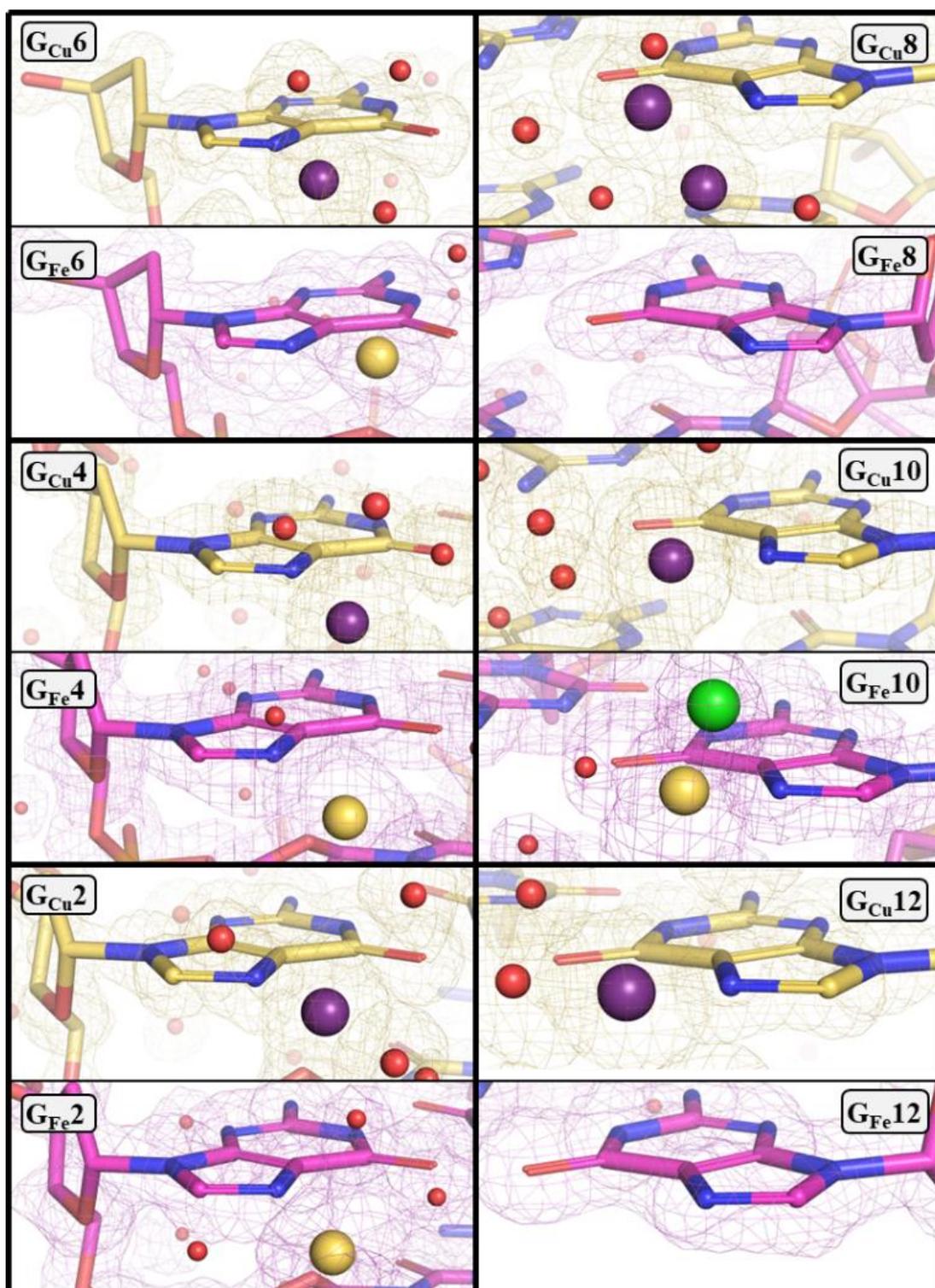


Figure 4.8: Comparison between all guanines present in the asymmetric unit for 8OEY and 8OEZ. Carbon atoms are shown as yellow and magenta in 8OEY and 8OEZ respectively. Oxygen, nitrogen, and phosphate atoms are shown as red, blue, and orange respectively. Copper ions are shown as purple spheres.

4.4. Discussion

Two novel crystal structures of the d(CGCGCG) sequence are presented in this work and are the result of a soaking experiment. **8OEY** shows the presence of six copper (II) ions sites bound to the duplex at the N7 positions of G_{Cu2}, G_{Cu4}, G_{Cu6}, G_{Cu8}, G_{Cu10}, and G_{Cu12}. All six were found to be partially occupying their position. Two of the Cu²⁺ were identified as 1.51 Å apart, which is a significantly shorter distance than copper-copper bond lengths found in the literature^{25,26}. For this reason, these two positions were modelled as occupied by the same copper. By comparison, 1D39, a structure of the d(CGCGCG) duplex in the presence of copper (II) ions has been previously reported²⁷. All six guanines were also partially coordinated. However, the native crystal used for this soaking experiment was grown in conditions lacking spermine and the Z-DNA duplex would have crystallised in the A polymorphic form¹⁷. A slightly different crystal packing might explain why one of the coppers bridge two guanines, while the same position in **8OEY** shows a high order of disorder and alternate positions (Figure 4.4). This shows the importance of highlighting crystal packing while discussing metal ions positions after a soaking experiment. However, this does not explain why each position is only partially occupied, as the soaking experiments of both polymorphic form (A and B) have shown partial occupancies at equivalent resolutions. In **8OEY**, all the copper (II) ions occupancies were reported as 51, 57, 48, 52, 51, 32, and 17 %. Yet, this phenomenon is not exclusive to copper coordination, as another soaking experiment of the B polymorphic form of Z-DNA has also led to structures containing partially occupied metal ions. Cr³⁺ ions were found to have partial occupancies of 35, 26 and 40 %¹⁸. In contrast, this is not the case for Mn²⁺ and Zn²⁺ ions, as these ions were fully coordinated at the N7 position of G6, and at the N7 position of G6 and G10, respectively²⁸. Opposingly, **8OEZ** shows the presence of four Fe²⁺ ions at 100, 100, 74, and 100 % occupancies, and were coordinated to the N7 positions of G_{Fe2}, G_{Fe4}, G_{Fe6}, and G_{Fe10} respectively. No Fe²⁺ ions were found to be coordinated at the N7 positions of G_{Fe8} and G_{Fe12}. Akin to Mn²⁺ and Zn²⁺ ions, and in contrast to Cu²⁺ and Cr³⁺ ions, three of the Fe²⁺ ions were found to be at full occupancy. Out of all the N7 positions occupied by metals reported in these studies, G6 appeared to be the preferred sites for metal binding, possibly because the positions can accommodate metals ions of various ionic radii. However, other factors must be at play for the other positions, like soaking time and concentration, as Zn²⁺ was found to bind less positions than Fe²⁺ with a smaller ionic radius²⁹.

These differences in coordination are significant, especially when studying the effects of oxidative damage. Differences in DNA oxidative damage have been observed to be specific to which metal catalyses the Fenton reaction and is often linked to differences in metal coordination to the DNA. For instance, in recent years, Burrows' group has characterised oxidation products depending on the metal present in the system and changes were attributed to the differences in binding affinity between copper and iron^{13,14}. While using a system using the nucleoside 2'-deoxyguanosine (dG) and the nucleotide 2'-deoxyguanosine-5'-monophosphate (dGMP) as a simple model to identify differences in interactions between Fe (II), Cu (II), Cr (III), and the base, Noblitt *et al.* have been able to observe differences in oxidative damage depending on the presence or absence of a phosphate group³⁰. The Fe(II)/H₂O₂ system gave higher yield of oxidative damage when reacting with the nucleotide, suggesting Fe (II) interacts with both N7 of the guanine base and the phosphate group, unlike Cu (II) which was found to interact with the N7 position on the guanine, and Cr (III), which was found to interact through the phosphate group³⁰. Additionally, capillary electrophoresis and Fourier transform infrared (FTIR) difference spectroscopic methods also confirmed Fe (II) binds to the phosphate groups and the N7 position of guanines³¹. However, extended X-ray absorption fine structure (EXAFS) of a solid Fe(II)-DNA complex provided evidence of five oxygen atoms and one nitrogen atom in the inner coordination sphere of the Fe(II)-DNA complex, supporting the idea that Fe (II) binds to the N7 position of a base, with no mention of binding through the phosphate groups³². Nonetheless, few structural studies have been done to confirm whether damage occurs where iron (II) is binding. For instance, Rai *et al.* used NMR to conduct a comprehensive study of iron (II) binding preferences following the observation that Fe (II)/H₂O₂ systems lead to sequence-specific strand-cleavage, especially at 5'-RTGR-3' sequences. And the resulting structure established that Fe (II) ions interact preferentially at the 5'-ATGA-3' site through the N7 position of the guanine. Another study by the same group later confirmed the correlation between preferential binding and enhanced cleavage at this position^{10,33}. However, no follow-up studies have been done to look at more systems such as A-DNA, Z-DNA or higher-order structures and this work remains an isolated example.

In **8OEY**, Fe (II) binds preferentially at the N7 position of guanines. No Fe (II) ions were found to bind to the phosphate backbone directly, however, water mediated contacts were identified between Fe²⁺ (1), Fe²⁺ (2) and Fe²⁺ (3) and phosphate groups from neighbouring Z-DNA hexamers. As such contacts were also identified in the d(CGCGCG)-Cu²⁺ structure however, it does not appear these interactions are typical of DNA-Fe²⁺ interactions as suggested by previous studies but are more of an artefact from crystal packing and the resulting proximity of other duplexes. Likewise, similar interactions have been identified in systems containing Ni (II)^{34,35}, Co (II)³⁶ and Zn (II)³⁷. Nonetheless, the major difference between the d(CGCGCG)-Cu²⁺ structure and the d(CGCGCG)-Fe²⁺ structure is the absence of metal centres at G8 and G12 in the later structure. The lack of phosphate groups at proximity however might explain the difference, as Cu (II) has the lowest relative metal affinity to the phosphate backbone and is known to bind preferentially to the N7 position of guanines³⁸.

In this work, the distances between Fe (II) and the N7 position of guanines have been reported as 2.35 Å, 2.26 Å, 2.35 Å and 2.51 Å. Similarly, the distances at the same position in the d(CGCGCG)-Cu²⁺ structure have been reported as 1.89 Å, 2.01 Å, 1.98 Å and 1.97 Å respectively. On average, Fe²⁺-N7 distances were found to be greater than Cu²⁺-N7 distances and are comparable to the distance (2.22±0.02 Å) reported by Bertoncini *et al.* in a DNA-Fe²⁺ system³². Additionally, a crystallographic study found that the average high-spin Fe (II)-N distance was 2.17 Å³⁹, which suggests the presence of high-spin Fe (II) bound to the DNA⁴⁰. However, a similar bond length between Fe (III) and N, along with a similar octahedral geometry for the hydration sphere between Fe (II) and Fe (III), might make it difficult to differentiate between Fe (III) ions and Fe (II) ions in the structure. The experiment was conducted in an oxygen-free environment; however, oxidation might still be possible during the transfer between the glove bag and the X-Ray diffractometer. Fortunately, Fe (II) oxidises at a slower rate at pH values close to the crystal system's pH in an aerobic environment, where the reaction is assumed to take hours, compared to oxidation at higher pH (> 7), where the reaction is believed to only take minutes¹⁶, and Fe (II) conversion to Fe (III) seems to be unlikely in the short time during the transfer to the X-Ray diffractometer.

Identifying Fe (II) metal centres is the first step in attempting in-crystal Fenton reaction. This work shows where potential damage hotspot could occur and could help understand where damage occurs as a function of structure in the crystalline environment. Interestingly, Burrows' group has reported, in the absence of reductants, a lower guanine conversion rate with the Fe(II)/H₂O₂ system¹³ than with the Cu(II)/H₂O₂ system¹⁴; this could be interesting to see if this difference can be explained by the lower distribution of binding sites in the d(CGCGCG)-Fe²⁺ structure compared to the d(CGCGCG)-Cu²⁺ structure.

4.5. References

1. Solomon, E. I., Decker, A. & Lehnert, N. Non-heme iron enzymes: Contrasts to heme catalysis. *Proc. Natl. Acad. Sci.* **100**, 3589–3594 (2003).
2. Nakae, Y., Stoward, P. J., Bessalov, I. A., Melamede, R. J. & Wallace, S. S. A new technique for the quantitative assessment of 8-oxoguanine in nuclear DNA as a marker of oxidative stress. Application to dystrophin-deficient DMD skeletal muscles. *Histochem. Cell Biol.* **124**, 335–345 (2005).
3. Walter, T., Pizarro, F., Abrams, S. A. & Boy, E. Bioavailability of elemental iron powder in white wheat bread. *Eur. J. Clin. Nutr.* **58**, 555–558 (2004).
4. Fenton, H. J. H. LXXIII.—Oxidation of tartaric acid in presence of iron. *J. Chem. Soc. Trans.* **65**, 899–910 (1894).
5. Haber, F., Weiss, J. & Pope, W. J. The catalytic decomposition of hydrogen peroxide by iron salts. *Proc. R. Soc. Lond. Ser. - Math. Phys. Sci.* **147**, 332–351 (1997).
6. Betteridge, D. J. What is oxidative stress? *Metab. - Clin. Exp.* **49**, 3–8 (2000).
7. Zuo, L., Zhou, T., Pannell, B. K., Ziegler, A. C. & Best, T. M. Biological and physiological role of reactive oxygen species – the good, the bad and the ugly. *Acta Physiol.* **214**, 329–348 (2015).
8. Brieger, K., Schiavone, S., Jr, F. J. M. & Krause, K.-H. Reactive oxygen species: from health to disease. *Swiss Med. Wkly.* **142**, 13659–13659 (2012).
9. Hofer, T. & Perry, G. Nucleic acid oxidative damage in Alzheimer’s disease—explained by the hepcidin-ferroportin neuronal iron overload hypothesis? *J. Trace Elem. Med. Biol.* **38**, 1–9 (2016).
10. Henle, E. S. *et al.* Sequence-specific DNA Cleavage by Fe²⁺-mediated Fenton Reactions Has Possible Biological Implications. *J. Biol. Chem.* **274**, 962–971 (1999).

11. Chevion, M. A site-specific mechanism for free radical induced biological damage: The essential role of redox-active transition metals. *Free Radic. Biol. Med.* **5**, 27–37 (1988).
12. Meneghini, R. Iron Homeostasis, Oxidative Stress, and DNA Damage. *Free Radic. Biol. Med.* **23**, 783–792 (1997).
13. Alshykhly, O. R., Fleming, A. M. & Burrows, C. J. 5-Carboxamido-5-formamido-2-iminohydantoin, in Addition to 8-oxo-7,8-Dihydroguanine, Is the Major Product of the Iron-Fenton or X-ray Radiation-Induced Oxidation of Guanine under Aerobic Reducing Conditions in Nucleoside and DNA Contexts. *J. Org. Chem.* **80**, 6996–7007 (2015).
14. Fleming, A. M., Muller, J. G., Ji, I. & Burrows, C. J. Characterization of 2'-deoxyguanosine oxidation products observed in the Fenton-like system Cu(II)/H₂O₂/reductant in nucleoside and oligodeoxynucleotide contexts. *Org. Biomol. Chem.* **9**, 3338–3348 (2011).
15. Rai, P., Cole, T. D., Wemmer, D. E. & Linn, S. Localization of Fe²⁺ at an RTGR sequence within a DNA duplex explains preferential cleavage by Fe²⁺ and H₂O₂. *J. Mol. Biol.* **312**, 1089–1101 (2001).
16. Santana-Casiano, J. M., González-Dávila, M. & Millero, F. J. Oxidation of Nanomolar Levels of Fe(II) with Oxygen in Natural Waters. *Environ. Sci. Technol.* **39**, 2073–2079 (2005).
17. Brzezinski, K. *et al.* High regularity of Z-DNA revealed by ultra high-resolution crystal structure at 0.55 Å. *Nucleic Acids Res.* **39**, 6238–6248 (2011).
18. Drozdal, P., Gilski, M., Kierzek, R., Lomozik, L. & Jaskolski, M. High-resolution crystal structure of Z-DNA in complex with Cr³⁺ cations. *JBIC J. Biol. Inorg. Chem.* **20**, 595–602 (2015).
19. McCoy, A. J. *et al.* Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674 (2007).

20. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2126–2132 (2004).
21. Afonine, P. V. *et al.* Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr. D Biol. Crystallogr.* **68**, 352–367 (2012).
22. Schrödinger, LLC. The PyMOL Molecular Graphics System, Version 1.8. (2015).
23. Zheng, G., Lu, X.-J. & Olson, W. K. Web 3DNA—a web server for the analysis, reconstruction, and visualization of three-dimensional nucleic-acid structures. *Nucleic Acids Res.* **37**, W240–W246 (2009).
24. Macrae, C. F. *et al.* Mercury 4.0: from visualization to analysis, design and prediction. *J. Appl. Crystallogr.* **53**, 226–235 (2020).
25. Bauschlicher, C. W., Walch, S. P. & Siegbahn, P. E. M. On the nature of the bonding in Cu₂. *J. Chem. Phys.* **76**, 6015–6017 (1982).
26. Hope, H. & Power, P. P. X-ray crystal structure of a late-transition-metal dialkylamide, [(CuNEt₂)₄]. *Inorg. Chem.* **23**, 936–937 (1984).
27. Kagawa, T. F., Geierstanger, B. H., Wang, A. H. & Ho, P. S. Covalent modification of guanine bases in double-stranded DNA. The 1.2-Å Z-DNA structure of d(CGCGCG) in the presence of CuCl₂. *J. Biol. Chem.* **266**, 20175–20184 (1991).
28. Drozdal, P., Gilski, M., Kierzek, R., Lomozik, L. & Jaskolski, M. Ultrahigh-resolution crystal structures of Z-DNA in complex with Mn²⁺ and Zn²⁺ ions. *Acta Crystallogr. D Biol. Crystallogr.* **69**, 1180–1190 (2013).
29. Shannon, R. D. Revised effective ionic radii and systematic studies of interatomic distances in halides and chalcogenides. *Acta Crystallogr. A* **32**, 751–767 (1976).
30. Noblitt, S. D., Huehls, A. M. & Morris, D. L. The role of metal ion binding in generating 8-hydroxy-2'-deoxyguanosine from the nucleoside 2'-deoxyguanosine and the nucleotide 2'-deoxyguanosine-5'-monophosphate. *J. Inorg. Biochem.* **101**, 536–542 (2007).

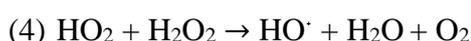
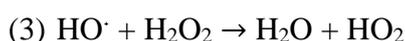
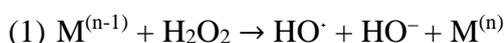
31. Ouameur, A. A., Arakawa, H., Ahmad, R., Naoui, M. & Tajmir-Riahi, H. a. A Comparative Study of Fe(II) and Fe(III) Interactions with DNA Duplex: Major and Minor Grooves Bindings. *DNA Cell Biol.* **24**, 394–401 (2005).
32. Bertoncini, C. R. A., Meneghini, R. & Tolentino, H. Nearest-neighbor nitrogen and oxygen distances in the iron(II)–DNA complex studied by extended X-ray absorption fine structure. *Spectrochim. Acta. A. Mol. Biomol. Spectrosc.* **77**, 908–910 (2010).
33. Rai, P., Wemmer, D. E. & Linn, S. Preferential binding and structural distortion by Fe²⁺ at RGGG-containing DNA sequences correlates with enhanced oxidative cleavage at such sequences. *Nucleic Acids Res.* **33**, 497–510 (2005).
34. Valls, N., Usón, I., Gouyette, C. & Subirana, J. A. A Cubic Arrangement of DNA Double Helices Based on Nickel–Guanine Interactions. *J. Am. Chem. Soc.* **126**, 7812–7816 (2004).
35. Abrescia, N. G., Huynh-Dinh, T. & Subirana, J. A. Nickel-guanine interactions in DNA: crystal structure of nickel-d[CGTGTACACG]₂. *JBIC J. Biol. Inorg. Chem.* **7**, 195–199 (2002).
36. Valls, N., Wright, G., Steiner, R. A., Murshudov, G. N. & Subirana, J. A. DNA variability in five crystal structures of d(CGCAATTGCG). *Acta Crystallogr. D Biol. Crystallogr.* **60**, 680–685 (2004).
37. Labiuk, S. L., Delbaere, L. T. J. & Lee, J. S. Cobalt(II), nickel(II) and zinc(II) do not bind to intra-helical N(7) guanine positions in the B-form crystal structure of d(GGCGCC). *JBIC J. Biol. Inorg. Chem.* **8**, 715–720 (2003).
38. Eichhorn, G. L., Butzow, J. J. & Shin, Y. A. Some effects of metal ions on DNA structure and genetic information transfer. *J. Biosci.* **8**, 527–535 (1985).
39. Zheng, H. *et al.* Data mining of iron(II) and iron(III) bond-valence parameters, and their relevance for macromolecular crystallography. *Acta Crystallogr. Sect. Struct. Biol.* **73**, 316–325 (2017).

40. Gütlich, P., Jung, J. & Goodwin, H. A. Spin Transitions in Iron(II) Complexes. in *Molecular Magnetism: From Molecular Assemblies to the Devices* 327–378 (1996).

Chapter 5 - Structural Analysis of Peroxide-Soaked DNA Crystals Containing Ordered Copper Binding Sites: Towards Understanding Oxidative Damage at the Atomic Scale

5.1. Introduction

Hydrogen peroxide (H_2O_2), a strong oxidant first discovered and reported as “oxygenated water”¹, has been described as a major redox metabolite of oxygen generated in multitude of biological processes. As mentioned in previous chapters, the hydrogen peroxide decomposition² (Scheme 5.1) can be catalysed by metal ions in the so-called Fenton reaction (Scheme 5.1.(1)) and generate hydroxyl radicals ($\text{HO}\cdot$), a highly reactive oxygen species (ROS) which can damage cellular components³.



Scheme 5.1: The reaction of copper with hydrogen peroxide

In normal circumstances, antioxidants keep ROS levels balanced. However, during oxidative stress, or when ROS production and the ability to neutralise ROS in a cell are unbalanced, many cellular components are damaged, and various disorders, including neurological disorders⁴, cancer, and cardiovascular diseases⁵ have been linked to this phenomenon. Nucleic acids are no exception to the rule and have been determined as being damaged easily in an oxidative stress environment. The presence of ROS in the cell can result in single strand breaks, double-strand breaks, the formation apurinic/apyrimidinic lesions and base modifications, and lead to mutations and diseases⁶.

The presence of metal ions and oxidative DNA damage and have frequently been linked. For instance, increased level of oxidative damage reflected by the presence of the oxidative stress marker product 8-oxo-7,8-dihydro-2'-deoxyguanine (oxoG) have been reported in the brains of individuals having suffered Alzheimer's diseases and have been associated to iron dysregulation⁷. And as discussed previously, binding preferences of metal ions have been associated to the type of oxidative damage found in nucleic acids, especially as hydroxyl radicals are short lived and a certain proximity between the reaction centre and where the damage occurs would be needed⁸⁻¹⁰.

In Chapter 4, a crystal system which could be soaked in both iron (II) chloride and copper (II) chloride was introduced. Both iron and copper catalyse the hydrogen peroxide decomposition through the Fenton reaction or a Fenton-like reaction respectively, however, the necessity to have anoxic conditions for the iron-soaked system made it impractical to try to soak iron-soaked crystals in hydrogen peroxide. Nevertheless, copper (II) can still act as an excellent model to develop a method to achieve in-crystal oxidative DNA damage.

This is not the first-time hydrogen peroxide soaking of crystals has been attempted, as this method has been attempted previously to study the active site of enzymes such as peroxide reductases¹¹, oxidoreductases¹² or superoxide dismutases¹³. However, soaking is not the only method to obtain ordered hydrogen peroxide in crystal systems, as the molecule could have been present before crystallisation. With this in consideration, a total of 81 crystallographic entries published on the Protein Data Bank (PDB) have reported hydrogen peroxide as a standalone ligand. Remarkably, none of the entries included nucleic acids. This highlights the lack of structural data needed to understand the interactions between nucleic acids, metal ions and oxidant leading to oxidative DNA damage.

In this Chapter, hydrogen peroxide soaking was attempted on copper-soaked crystals. A long soak time was chosen after taking crystal size into consideration as well as average reaction times for Cu (II)/H₂O₂ systems in solution¹⁴. However, this led to three different structures,

8OE7, 8OE8 and **8OE9**, two of which have hydrogen peroxide molecules trapped, and will be discussed here. Additionally, to confirm whether copper (II) ions slowly disappear as the reaction take place or if copper-soaking simply did not result in all guanines occupied by copper (II) ions, three data sets were collected from a single crystal on three different occasions during the experiment. This was possible by collecting data from the crystal at room temperature using a humidity control device, to stabilise the crystal and prevent desiccation. The copper (II) ions disappearance in **8OEA, 8OEB** and **8OEC** will be discussed in conjunction with the result from **8OE7, 8OE8** and **8OE9**.

5.2. Material and methods

5.2.1. DNA preparation

The oligonucleotide d(CGCGCG) was purchased from Eurogentec as a solid purified by RP-HPLC. DNA concentration was determined from the value of the absorbance at $\lambda = 260$ nm, using the molar extinction coefficient provided by the manufacturer. 1.5 mM solution of the DNA was annealed at 358 K for two minutes then allowed to cool slowly to room temperature.

5.2.2. Crystallisation of d(CGCGCG)

Single crystals to be used for soaking in Fenton reagents were grown at 291 K by the sitting-drop vapor diffusion method using the same protocol used in Chapter 4¹⁵ and described by Drozdal *et al.* The drops contained 2 μ L of 1.5 mM oligonucleotides and 2 μ L of precipitating solution containing 10 % (v/v) 2-methyl-2,4-pentanediol (MPD), 40 mM sodium cacodylate, pH 6.0, 80 mM KCl, 12 mM NaCl and 12 mM spermine tetrachloride. The drops were equilibrated against 0.5 ml of 35 % (v/v) MPD. The crystals appeared within 48 hours.

5.2.3. Time-dependent in-crystal Fenton reaction and flash-freezing

Dozens of d(CGCGCG) crystals were transferred to a drop containing 2 μ L of precipitating solution and 2 μ L of copper (II) chloride and left to soak for 24 hours. The crystals

were then transferred to a drop containing 2 μL of precipitating solution and 2 μL of 2 % hydrogen peroxide. The crystals were removed at different time intervals, ranging from 5 minutes to 1 hour, back-soaked for 30 seconds in mother liquor, and flash-cooled in liquid nitrogen to be sent to Diamond Light Source for collection.

5.2.4. Synchrotron data collection and processing

Data from flash-cooled crystals of approximate dimensions 200 x 100 x 100 μm were collected at 100 K with 0.004 s frames on beam line I03 at Diamond Light Source Ltd using radiation with a wavelength of 0.9763 \AA . For each dataset, 3600 frames were collected with an oscillation angle of 0.10° to give 360° of data. The data were integrated and scaled using xia2¹⁶ with Dials. Out of the 17 crystals collected, 13 gave sufficient diffraction. The structures were solved in $P2_12_12_1$ by molecular replacement using Phaser¹⁷ with the structure **1I0T** from the PDB as the starting model, except for two, which had to be solved in P1. Three structures were selected to be presented in this work for their structural differences.

Data were collected from a crystal soaked in hydrogen peroxide for 5 minutes and gave 11347 unique reflections after being integrated and scaled. The structure was initially solved in $P2_12_12_1$ by molecular replacement, however, high R-factors after multiple refinement cycles suggested this might not have been the appropriate strategy. Consequently, the data were integrated and scaled again, this time in P1, giving 24395 unique reflections. After molecular replacement, the model was updated using Coot¹⁸ and refined using phenix.refine¹⁹. To confirm that the model was satisfactory and that the observed data and the calculated data from the model were in agreement, the R_{free} and the R_{work} were calculated after each round of refinement. R-factors can be made arbitrarily low by overfitting the data during refinement. Consequently, five percent of the reflections were reserved for the R_{free} set and not used during refinement. The R_{free} and R_{work} observed after the final round of refinement for this model were 0.2332 of 0.1994 respectively. These values can be compared relatively to X-ray structures of similar resolution after deposition and validation in the Protein Data Bank. The data, and final coordinates, were deposited in the Protein Data Bank with ID **8OE7**.

A second data set was collected from a crystal soaked in hydrogen peroxide for 1 hour. The initial structure was also solved in $P2_12_12_1$ by molecular replacement, however, high R-factors after multiple refinement cycles were also observed, and the data were integrated and scaled in $P1$, giving 21952 unique reflections. After molecular replacement, the model was updated using Coot¹⁸ and refined using phenix.refine¹⁹ to give a final R_{free} of 0.2089 and R_{work} of 0.1691. Five percent of reflections were reserved for the R_{free} set. The data, and final coordinates, were deposited in the Protein Data Bank with ID **8OE8**.

The third dataset was collected from a crystal soaked in hydrogen peroxide for 1 hour and gave 11394 unique reflections after being integrated and scaled. The structure was solved in $P2_12_12_1$ by molecular replacement. The model was updated using Coot¹⁸ and refined using phenix.refine¹⁹ to give a final R_{free} of 0.1779 and R_{work} of 0.1551. Five percent of reflections were reserved for the R_{free} set. The data, and final coordinates, were deposited in the Protein Data Bank with ID **8OE9**. The data collection and refinement statistics of all three datasets for can be found in Table 5.1.

Table 5.1: Synchrotron data collection and refinement statistics for 8OE7, 8OE8 and 8OE9. Outer shell statistics shown in parentheses.

PDB ID	8OE7	8OE8	8OE9
Data collection			
Space group	P1	P1	P2 ₁ 2 ₁ 2 ₁
Cell dimensions			
<i>a</i> , <i>b</i> , <i>c</i> , (Å)	18.28, 30.74, 42.5	18.35, 30.67, 42.86	18.340, 30.710, 42.630
α , β , γ (°)	90.01, 89.95, 90.01	89.98, 90.06, 89.94	
Resolution, (Å)	30.74 - 1.25 (1.27 - 1.25)	42.86 - 1.30 (1.32 - 1.30)	24.92 - 1.04 (1.06 - 1.04)
R _{merge}	0.043 (0.226)	0.051 (0.431)	0.186 (1.312)
R _{meas}	0.050 (0.264)	0.060 (0.508)	0.189 (1.539)
R _{pim}	0.027 (0.136)	0.032 (0.267)	0.060 (0.772)
Total number of observations	87060	79075	110564
Total number of unique observations	24395	21952	11394
CC _{1/2}	1.000 (0.940)	0.998 (0.789)	0.996 (0.357)
I/ σ I	19.7 (3.3)	19.9 (3.0)	6.7 (1.1)
Completeness, (%)	95.1 (89.3)	95.4 (89.3)	94.0 (55.1)
Multiplicity	3.6 (3.7)	3.6 (3.6)	9.7 (3.6)
Refinement			
Resolution, (Å)	24.920 - 1.250 (1.295 - 1.250)	21.430 - 1.300 (1.346 - 1.300)	24.918 - 1.040 (1.077 - 1.040)
No. Reflections	24320 (2345)	21922 (2090)	11327 (716)
R _{free} /R _{work}	0.2332/0.1994	0.2089/0.1691	0.1779/0.1551
rmsd			
Bond lengths, (Å)	0.010	0.010	0.012
Bond angles, (°)	1.32	1.32	1.34

5.2.5. In-house multi dataset collection at room-temperature from a peroxide-soaked crystal as a function of time

A d(CGCGCG) crystal was transferred to a drop containing 2 μ L of precipitating solution and 2 μ L of 20 mM copper (II) chloride and left to soak for 24 hours (Figure 5.1). The crystal was mounted on a Micromesh from MiTeGen and placed on the goniometer of the in-house XtaLAB Synergy-S X-ray diffractometer. The crystal was kept hydrated at a relative humidity of 99 % using a HC-Lab Humidity Controller and a data set was collected. 872 frames were collected from the crystal. The data were integrated and scaled using the CrysAlisPRO software. The structure was initially solved in P2₁2₁2₁ by molecular replacement using Phaser¹⁷ with the structure **1I0T** from the PDB as the starting model. However, high R-factors after multiple refinement cycles suggested this might not have been the appropriate strategy. Consequently, the data were integrated and scaled again, this time in P2₁, giving 4161 unique reflections. Data quality was assessed using phenix.xtriage and the intensity statistics suggested twinning. The model was updated using Coot¹⁸ and refined using phenix.refine¹⁹. Twinned

refinement was carried out using the h, -k, -l operator to give a final R_{free} of 0.2276 and R_{work} of 0.1786. Five percent of reflections were reserved for the R_{free} set. The data, and final coordinates, were deposited in the Protein Data Bank with ID **8OEA**.

The loop was then unmounted and transferred to a drop containing 2 μL of precipitating solution and 2 % of hydrogen peroxide and left to soak for 30 minutes (Figure 5.1). The same crystal was then mounted again on a Micromesh and another data set was collected. 326 frames were collected from the crystal. The data were integrated and scaled using the CrysAlisPRO software, giving 2494 unique reflections. The structure was solved in $P2_12_12_1$ by molecular replacement using Phaser¹⁷ with the structure **1I0T** from the PDB as the starting model. The model was updated using Coot¹⁸ and refined using phenix.refine¹⁹ to give a final R_{free} of 0.1818 and R_{work} of 0.1737. Five percent of reflections were reserved for the R_{free} set. The data, and final coordinates, were deposited in the Protein Data Bank with ID **8OEB**.

The same soaking step was repeated one more time and a third data set was collected from the same crystal (Figure 5.1). 250 frames were collected from the crystal. The data were integrated and scaled using the CrysAlisPRO software, giving 1612 unique reflections. The structure was solved in $P2_12_12_1$ by molecular replacement using Phaser¹⁷ with the structure **1I0T** from the PDB as the starting model. The model was updated using Coot¹⁸ and refined using phenix.refine¹⁹ to give a final R_{free} of 0.1880 and R_{work} of 0.1467. Five percent of reflections were reserved for the R_{free} set. The data, and final coordinates, were deposited in the Protein Data Bank with ID **8OEC**. The data collection and refinement statistics of all three datasets for can be found in Table 5.2.

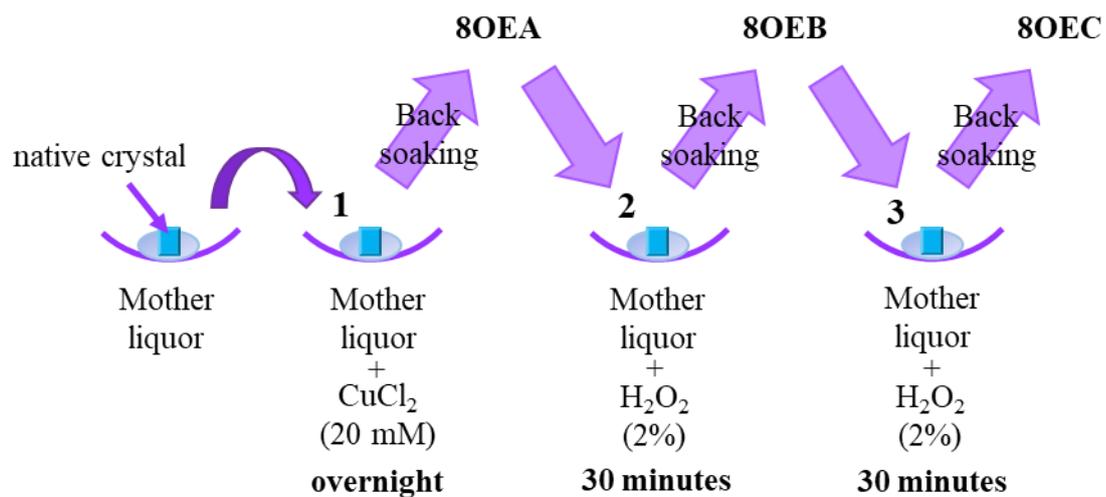


Figure 5.1: Diagram to illustrate when all three data sets were collected from the crystal during the peroxide-soaked experiment. Step 1 involved soaking the crystal in a drop containing a 20 mM copper (II) chloride mixed with the crystallisation condition. Step 2 and step 3 involved soaking the crystal in a drop containing a 2 % H₂O₂ mixed with the crystallisation condition. The crystal was rinsed each time for thirty seconds prior to collection.

Table 5.2: In-house data collection and refinement statistics for 8OEA, 8OEB and 8OEC. Outer shell statistics shown in parentheses.

PDB ID	8OEA	8OEB	8OEC
Data collection			
Space group	P2 ₁	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁
Cell dimensions			
<i>a, b, c</i> , (Å)	18.09, 30.65, 43.00	18.31, 30.78, 43.16	18.36, 30.78, 43.10
Resolution, (Å)	18.09 - 1.50 (1.54 - 1.50)	18.31 - 1.80 (1.88 - 1.80)	17.65 - 2.10 (2.25 - 2.10)
R _{merge}	0.078 (0.327)	0.059 (0.208)	0.069 (0.161)
R _{meas}	0.086 (0.390)	0.073 (0.270)	0.088 (0.205)
R _{pim}	0.036 (0.209)	0.042 (0.170)	0.054 (0.125)
Total number of observations	35360	12004	6194
Total number of unique observations	4161	2494	1612
CC _{1/2}	1.000 (0.885)	0.999 (0.952)	0.999 (0.982)
I/σI	14.6 (3.3)	13.5 (5.4)	10.7 (6.7)
Completeness, (%)	99.9 (99.8)	99.6 (100.0)	99.5 (100.0)
Multiplicity	8.5 (5.6)	4.8 (3.8)	3.8 (4.2)
Refinement			
Resolution, (Å)	18.091 - 1.500 (1.554 - 1.500)	17.670 - 1.800 (1.864 - 1.800)	17.654 - 2.100 (2.175 - 2.100)
No. Reflections	7645 (754)	2461 (234)	1586 (151)
R _{free} /R _{work}	0.2276/0.1786	0.1818/0.1737	0.1880/0.1467
rmsd			
Bond lengths, (Å)	0.004	0.010	0.004
Bond angles, (°)	0.64	1.83	1.67

5.2.6. Structural and packing analysis.

All figures were created using PyMOL²⁰ unless specified otherwise. Dinucleotide step parameters were calculated using Web 3DNA 2.0²¹. CCP4 maps were generated from the final mtz file, including the anomalous map when available, using the phenix.mtz2map program. Maps were visualised in PyMOL and displayed using the mesh function. *2mFo-DFc* maps were contoured at 1σ unless specified otherwise. Anomalous maps were contoured at 2σ unless specified otherwise. Packing diagrams were generated using Mercury²².

5.3. Results

5.3.1. Peroxide-soaked d(CGCGCG) crystals: 100 K collection.

5.3.1.1 Overall structure and helical parameters analysis

8OE7, **8OE8** and **8OE9** were obtained from crystals of the DNA hexamer sequence d(CGCGCG). All three were soaked in the same drop containing hydrogen peroxide, flash-cooled and collected at 100 K.

8OE7, the first peroxide-soaked structure, was obtained from a native crystal of the DNA sequence d(CGCGCG) grown in the same conditions as **8OEX**, which was previously soaked in copper (II) chloride for 24 hours. After soaking the crystal in hydrogen peroxide for five minutes, the crystal was flash-frozen and sent for collection. The crystal diffracted to 1.25 Å, a resolution close to the resolution of the other structures collected during this experiment, and the diffraction quality did not deteriorate as much as expected. However, the shift to the triclinic space group P1 and a high R_{Free} could indicate that the soaking modified the lattice, although crystal packing is consistent with the polymorphic form B of d(CGCGCG) Z-DNA structures and does not indicate extensive changes in the crystal lattices. The asymmetric unit contains eight strands for a total of 48 nucleotides and forms four duplexes. No spermine molecules were found in the asymmetric unit. The nucleotides from each chain (A, B, C, D, E, F, G, H) are labelled from $C_{\text{Chain}1}$ to $G_{\text{Chain}6}$ in the 5'-3' direction. For clarity, dinucleotide parameters were calculated separately for each duplex in the structure (A/B, C/D, E/F and G/H). The structural parameters classify each duplex as a left-handed Z-form structure. Disorder is observed at the $G_{A4}-C_{A5}$, $G_{B2}-C_{B3}$, $G_{C4}-C_{C5}$, $G_{D4}-C_{D5}$, $G_{E4}-C_{E5}$ and $G_{F2}-C_{F3}$ phosphate linkages. The alternative conformations are referred to as 1 and 2, and have occupancies of 62/38, 59/41, 67/33, 51/49, 64/36 and 47/53 % respectively. Details of the local base-pair parameters can be found in the Supplementary Information.

8OE8, the second peroxide-soaked structure, was obtained from a crystal grown and soaked in similar conditions as **8OE7**, except for length of time this crystal was left in hydrogen peroxide conditions. This crystal was soaked in hydrogen peroxide for an hour. The crystal diffracted to 1.30 Å. The asymmetric unit contains eight strands for a total of 48 nucleotides

and forms four duplexes. A spermine molecule was also found to interact with each Z-DNA duplexes in the crystal lattice, to a total of four molecules in the asymmetric unit; but only the ordered part of the molecule was modelled, as the rest was found to be highly disordered and could not be modelled in electron density. The nucleotides from each chain (A', B', C', D', E', F', G', H') are labelled from C_{Chain1} to G_{Chain6} in the 5'-3' direction. As **8OE7**, dinucleotide parameters were calculated separately for each duplex in the structure (A'/B', C'/D', E'/F' and G'/H'). Soaking did not change the overall structure parameters and each duplex is still classified as a left-handed Z-form DNA. Disorder is observed at the G_A'4-C_A'5, G_B'2-C_B'3, G_B'4-C_B'5, G_C'4-C_C'5, G_D'2-C_D'3, G_D'4-C_D'5, G_E'4-C_E'5, G_F'2-C_F'3, G_F'4-C_F'5, G_G'4-C_G'5, G_H'2-C_H'3 and G_H'4-C_H'5, phosphate linkages. The alternative conformations are referred to as 3 and 4, and have occupancies of 71/29, 24/76, 74/26, 67/33, 27/73, 63/37, 72/38, 30/70, 61/39, 71/29, 26/74 and 62/38 % respectively. Details of the local base-pair parameters can be found in the Supplementary Information. Equivalently to **8OE7**, the structure belongs to the space group P1, which could also indicate a compromised crystal lattice. However, crystal packing is still consistent with the polymorphic form B of d(CGCGCG) Z-DNA structures.

8OE9, the third peroxide-soaked structure, was obtained from a crystal grown and soaked in identical conditions as **8OE8**. The crystal diffracted to 1.04 Å. The asymmetric unit contains two strands for a total of 12 nucleotides and forms four duplexes. A spermine molecule was also found to interact with the two Z-DNA duplexes in the crystal lattice, in a similar way to the other Z-DNA structures presenting the same crystal packing. The nucleotides from each chain (A'' and B'') are labelled from C_{Chain1} to G_{Chain6} in the 5'-3' direction. The structural parameters classify the duplex as a left-handed Z-form structure. Disorder is observed at the G_B''2-C_B''3 phosphate linkage, as well as at the C_B''1 5'-OH group. The alternative conformations are referred to as 5 and 6 and have occupancies of 47/53 and 63/37 % respectively. Details of the local base-pair parameters can be found in the Supplementary Information. The structure belongs to the space group P2₁2₁2₁, and crystal packing is consistent with the polymorphic form B of d(CGCGCG) Z-DNA structures. Globally, none of the sugar pucker pseudorotation angles in **8OE7**, **8OE8** and **8OE9** differ significantly from the ones reported in **8OEX** and **8OEY**. Details of the sugar pucker parameters can be found in Table 5.3.

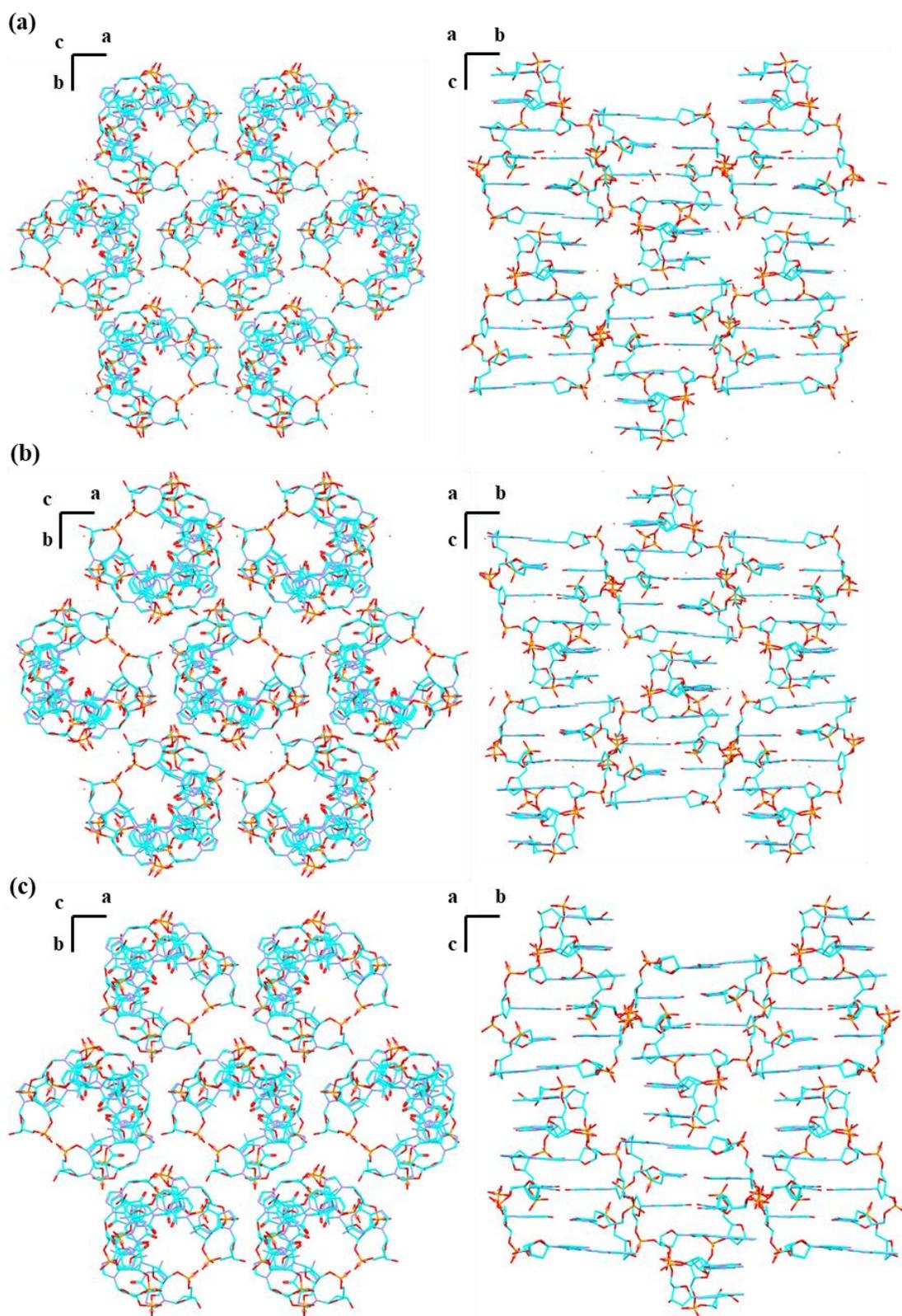


Figure 5.2: a. Packing diagram of the 8OE7 structure. b. Packing diagram of the 8OE8 structure. c. Packing diagram of the 8OE9 structure with views along the c-axes (left) and the a-axis (right).

Table 5.3: Sugar pucker pseudorotation angles and sugar pucker type for the oligonucleotide d(CGCGCG) in 8OE7, 8OE8 and 8OE9. Only chain A and chain B of each structure have been represented here. The sugar pucker type for the oligonucleotide d(CGCGCG) in 8OEX and 8OEY have been added as comparison. Chain C, D, E, F, G and H of structures 8OE7 and 8OE8 can be found in the Supplementary Information.

	Base	8OE7		8OE8		8OE9		8OEX	8OEY
		Phase angle of pseudorotation (P), °	Sugar Pucker	Phase angle of pseudorotation (P), °	Sugar Pucker	Phase angle of pseudorotation (P), °	Sugar Pucker	Sugar Pucker	Sugar Pucker
CHAIN A	5'-C	154.5	C2'-endo	152.3	C2'-endo	149.2	C2'-endo	C2'-endo	C2'-endo
	G	41.9	C4'-exo	42.5	C4'-exo	32.8	C3'-endo	C3'-endo	C3'-endo
	C	168.5	C2'-endo	166.6	C2'-endo	163.5	C2'-endo	C2'-endo	C2'-endo
	G	40.4	C4'-exo	45.0	C4'-exo	41.7	C4'-exo	C3'-endo	C4'-exo
	C	145.7	C2'-endo	143.1	C1'-exo	140.2	C1'-exo	C1'-exo	C1'-exo
	G-3'	168.7	C2'-endo	167.4	C2'-endo	169.2	C2'-endo	C2'-endo	C2'-endo
CHAIN B	5'-C	160.2	C2'-endo	152.7	C2'-endo	149.5	C2'-endo	C2'-endo	C2'-endo
	G	44.2	C4'-exo	44.5	C4'-exo	33.1	C3'-endo	C4'-exo	C3'-endo
	C	157.2	C2'-endo	167.8	C2'-endo	162.8	C2'-endo	C2'-endo	C2'-endo
	G	35.2	C3'-endo	38.1	C4'-exo	35.0	C3'-endo	C3'-endo	C3'-endo
	C	155.5	C2'-endo	151.6	C2'-endo	154.4	C2'-endo	C2'-endo	C2'-endo
	G-3'	167.6	C2'-endo	167.4	C2'-endo	168.0	C2'-endo	C2'-endo	C2'-endo

5.3.1.2 Coordination of the Cu²⁺ ions and H₂O₂ molecules

In **8OE7**, a total of twenty copper (II) ions were found after the native crystal was soaked in copper (II) chloride followed by soaking in hydrogen peroxide for five minutes. The Cu²⁺ ions were added based on the electron density present and the distance from the N7 position of the guanine. All Cu²⁺ ions were coordinated to the N7 position of guanines. Five Cu²⁺ ions were found coordinated to each duplex. And after occupancy refinement, a mix of partial and total occupancies were attributed to all the Cu²⁺ ions as well as partial coordination spheres. Each Cu²⁺ ion was named following this convention: Cu²⁺ (G_{Chain} 'Position') for clarity. Details of copper coordination are available in Table 5.4. Only the copper (II) coordinated to hydrogen peroxide molecules are listed.

Table 5.4: Copper (II) ions parameters from 8OE7, 8OE8 and 8OE9 (distance in Å)

PDB ID	Metal name	Occupancy metal	Distance to guanine N7
8OE7	Cu ²⁺ (GA4)	1.00	2.03
	Cu ²⁺ (GA4)	1.00	2.18
	Cu ²⁺ (GC2)	1.00	2.14
	Cu ²⁺ (GC4)	1.00	2.03
	Cu ²⁺ (GC4)	1.00	2.03
	Cu ²⁺ (GG4)	1.00	1.87
	Cu ²⁺ (GG6)	0.90	2.09
8OE8	Cu ²⁺ (GC2)	0.34	2.38

Interestingly, the coordination spheres of six of these Cu²⁺ ions, Cu²⁺ (GA4), Cu²⁺ (GA4), Cu²⁺ (GC2), Cu²⁺ (GC4), Cu²⁺ (GG4), Cu²⁺ (GG6) and Cu²⁺ (GC2) included hydrogen peroxide molecules. Cu²⁺ (GC4) was the only Cu²⁺ ion to have two hydrogen peroxide coordinated, Oxy (4) and Oxy (5). Details of hydrogen peroxide coordination are available in Table 5.5.

In **8OE8**, a total of eleven copper (II) ions were found after the native crystal was soaked in copper (II) chloride followed by soaking in hydrogen peroxide for an hour. The Cu²⁺ ions were added based on the electron density present and the distance from the N7 position of the guanine. All Cu²⁺ ions were coordinated to the N7 position of guanines. Three Cu²⁺ ions were found coordinated per duplex in C'/D', E'/F' and G'/H'. Two Cu²⁺ ions were found coordinated to duplex A'/B'. After occupancy refinement, a mix of partial and total occupancies were attributed to all the Cu²⁺ ions as well as partial coordination spheres. Details of copper coordination are available in Table 5.3. A single hydrogen peroxide molecule, H₂O₂ (1'), was also found coordinated to Cu²⁺ (GC2). Details of hydrogen peroxide coordination are available in Table 5.4.

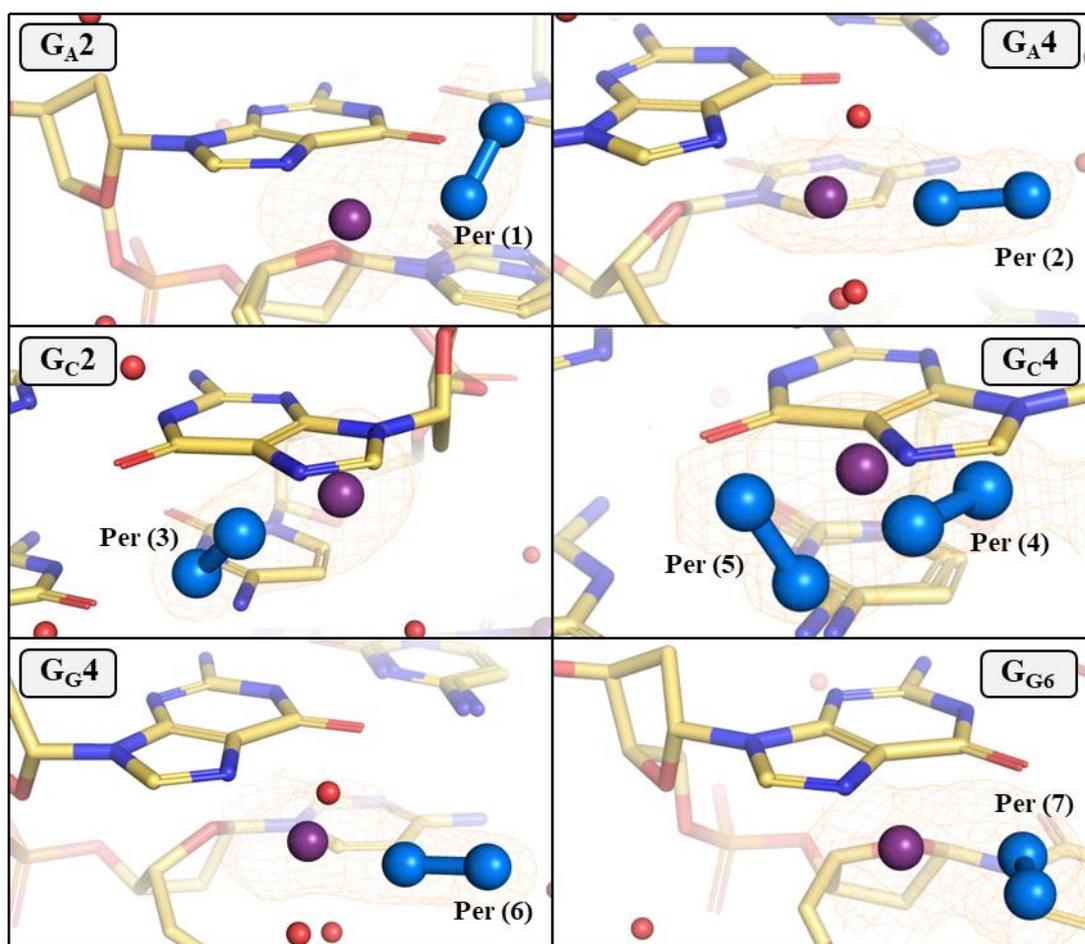


Figure 5.3: Close-up of guanines with at least one hydrogen peroxide present in the coordination sphere of its copper (II) ion. Carbon, oxygen, nitrogen, and phosphate atoms are shown as yellow, red, blue, and orange respectively. Copper (II) ions are represented as purple spheres. Hydrogen peroxide molecules are represented as blue spheres. The $2mF_o-DF_c$ electron density map is contoured at the 1.0σ level (orange).

Table 5.5: Hydrogen peroxide parameters from 8OE7 and 8OE8 (distance in Å)

PDB ID	H ₂ O ₂ Name	Occupancy H ₂ O ₂	Distance to Cu ²⁺	Metal name
8OE7	Per (1)	0.80	1.81	Cu ²⁺ (GA2)
	Per (2)	1.00	1.91	Cu ²⁺ (GA4)
	Per (3)	0.97	1.81	Cu ²⁺ (GC2)
	Per (4)	0.69	2.03	Cu ²⁺ (GC4)
	Per (5)	1.00	1.75 (O ₁); 1.73 (O ₂)	Cu ²⁺ (GC4)
	Per (6)	0.95	1.72	Cu ²⁺ (GG4)
	Per (7)	1.00	1.87	Cu ²⁺ (GG6)
8OE8	Per (1')	1.00	2.02	Cu ²⁺ (GC'2)

However, no copper (II) ions were found in **8OE9** after the native crystal was soaked in copper (II) chloride followed by soaking in hydrogen peroxide for an hour, unlike **8OE8**. Interestingly, these three crystals were treated in the same conditions but still showed widely different results. After copper (II) chloride soaking, a copper bound to each guanine in the asymmetric unit was to be expected in each crystal. However, after hydrogen peroxide soaking, some of the copper (II) ions were found to be missing in all three structures, especially in **8OE9** where none remained. And although these crystals were grown in the same conditions and were soaked in the same drops, the possibility that the copper (II) chloride soaking step was not performed long enough for all positions to be occupied remains. To determine if the absence of copper (II) ions was due to copper-soaking or peroxide-soaking, a single crystal was used to collect multiple data sets at different points of the soaking experiment.

5.3.2. Peroxide-soaked d(CGCGCG) crystals: room temperature collection.

5.3.2.1 Overall structure and helical parameters analysis

8OEA, **8OEB** and **8OEC** were collected from that same crystal of the DNA hexamer sequence d(CGCGCG). **8OEA** was obtained from a native crystal of the DNA sequence d(CGCGCG) grown in the same conditions as all the other structures in this Chapter and soaked in copper (II) chloride. This collection was performed at room temperature using a HC-Lab Humidity Controller to preserve the crystal and confirm the presence of copper (II) ions before the introduction of hydrogen peroxide into the system. The crystal diffracted to 1.50 Å. The asymmetric unit contains four strands for a total of 24 nucleotides and forms two duplexes. No spermine molecules were found in the asymmetric unit. The nucleotides from each chain (a, b, c, d) are labelled from C_{Chain1} to G_{Chain6} in the 5'-3' direction. Dinucleotide parameters were calculated separately for each duplex in the structure (a/b and c/d). The structural parameters classify each duplex as a left-handed Z-form structure. Disorder is observed at the G_{a4}-C_{a5}, G_{b2}-C_{b3} and G_{b4}-C_{b5} phosphate linkages. The alternative conformations are referred to as a and b and have occupancies of 22/78, 43/67 and 13/87 respectively. Details of the local base-pair parameters can be found in the Supplementary Information. The structure belongs to the space group P2₁, and crystal packing is consistent with the polymorphic form B of d(CGCGCG) Z-DNA structures.

8OEB was obtained after the same crystal used to collect **8OEA** on the in-house source was transferred to a drop containing hydrogen peroxide and left to soak for thirty minutes. This collection was performed at room temperature using a humidity controller to preserve the crystal for another soak and collection. The crystal did not diffract as much as **8OEA** and diffracted to 1.80 Å. The asymmetric unit contains two strands for a total of twelve nucleotides and forms a duplex. A spermine molecule was also found to interact in the major groove with two Z-DNA duplexes in the crystal lattice. The nucleotides from each chain (a' and b') are labelled from C_{Chain1} to G_{Chain6} in the 5'-3' direction. The structural parameters classify the duplex as a left-handed Z-form structure. Disorder is observed at the G_b'2-C_b'3 phosphate linkage. The alternative conformations are referred to as c and d and have occupancies of 36/64 %. Details of the local base-pair parameters can be found in the Supplementary Information. The structure belongs to the space group P2₁2₁2₁, and crystal packing is consistent with the polymorphic form B of d(CGCGCG) Z-DNA structures.

8OEC was obtained after the same crystal used to collect **8OEA** and **8OEB** was transferred one more time to a drop containing hydrogen peroxide and left to soak for another thirty minutes. This collection was still performed at room temperature using a humidity controller for consistency between the data sets. This time, the crystal diffracted to 2.10 Å. The asymmetric unit contains two strands for a total of twelve nucleotides and forms a duplex. A spermine molecule was also found to interact in the major groove with two Z-DNA duplexes in the crystal lattice. The nucleotides from each chain (a'' and b'') are labelled from C_{Chain1} to G_{Chain6} in the 5'-3' direction. The structural parameters classify the duplex as a left-handed Z-form structure. Disorder is observed at the G_b'2-C_b'3 phosphate linkage. The alternative conformations are referred to as e and f and have occupancies of 57/43 %. Details of the local base-pair parameters can be found in the Supplementary Information. The structure belongs to the space group P2₁2₁2₁, and crystal packing is consistent with the polymorphic form B of d(CGCGCG) Z-DNA structures.

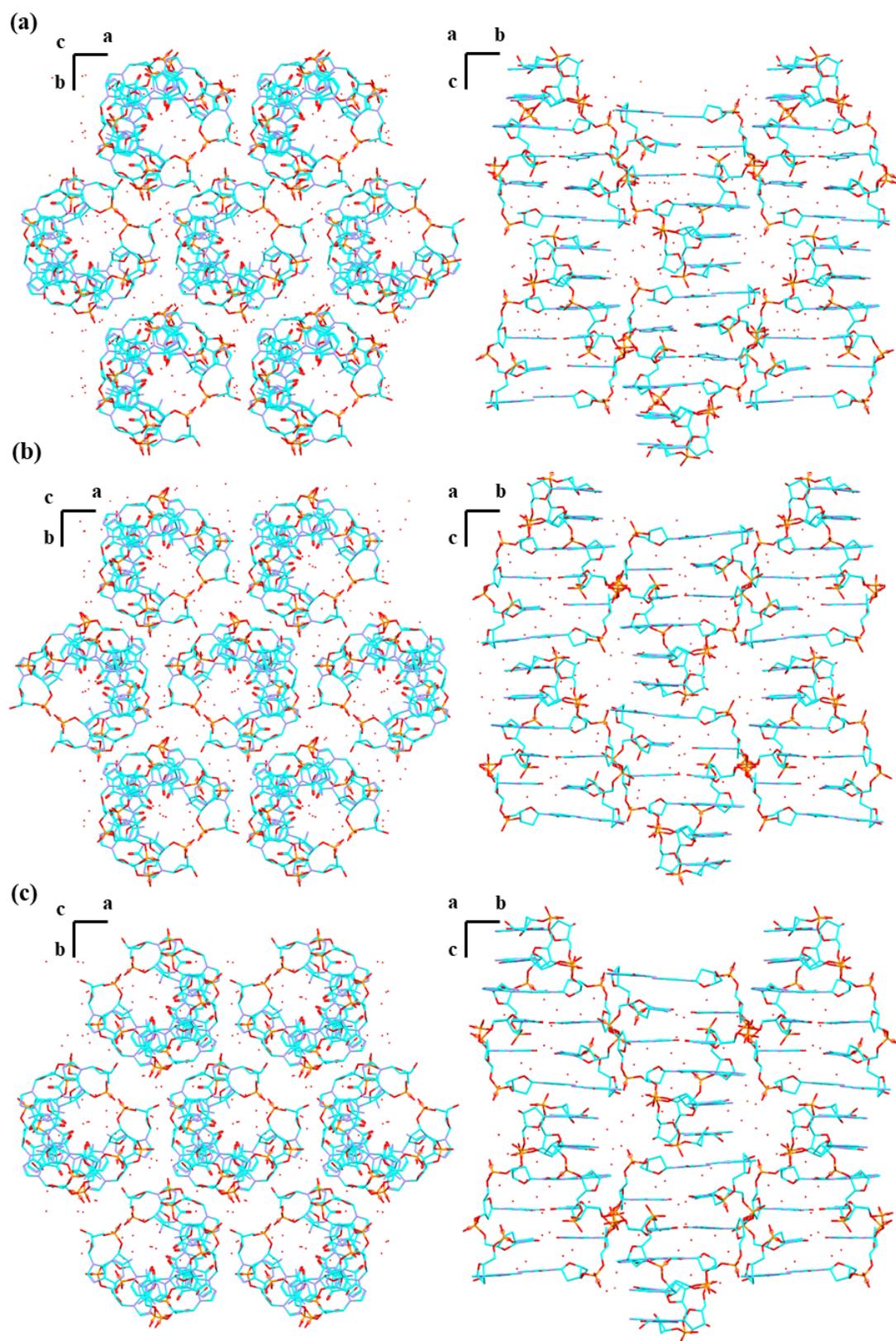


Figure 5.4: a. Packing diagram of the 8OEA structure. b. Packing diagram of the 8OEB structure. c. Packing diagram of the 8OEC structure with views along the c-axes (left) and the a-axis (right).

Interestingly, one of the sugar pucker pseudorotation angles in **8OEA** differed significantly from all the other ones reported in, **8OEB**, **8OEC**, **8OEX** and **8OEY**. The sugar pucker at G_b2 were classified as O4'-endo. Details of the sugar pucker parameters can be found in Table 5.6.

Table 5.6: Sugar pucker pseudorotation angles and sugar pucker type for the oligonucleotide d(CGCGCG) in 8OEA, 8OEB and 8OEC. Only chain a and chain b of each structure have been represented here. The sugar pucker type for the oligonucleotide d(CGCGCG) in 8OEX and 8OEY have been added as comparison. Chain c and d of structure 8OEA can be found in the Supplementary Information.

	Base	8OEA		8OEB		8OEC		8OEX	8OEY
		Phase angle of pseudorotation (P), °	Sugar Pucker	Phase angle of pseudorotation (P), °	Sugar Pucker	Phase angle of pseudorotation (P), °	Sugar Pucker	Sugar Pucker	Sugar Pucker
C H A I N a	5'-C	137.4	C1'-exo	169.2	C2'-endo	148.9	C2'-endo	C2'-endo	C2'-endo
	G	49.8	C4'-exo	43.4	C4'-exo	45.8	C4'-exo	C3'-endo	C3'-endo
	C	157.6	C2'-endo	164.9	C2'-endo	162.9	C2'-endo	C2'-endo	C2'-endo
	G	47.5	C4'-exo	60.0	C4'-exo	43.1	C4'-exo	C3'-endo	C4'-exo
	C	149.3	C2'-endo	150.7	C2'-endo	147.6	C2'-endo	C1'-exo	C1'-exo
	G-3'	169.2	C2'-endo	171.7	C2'-endo	162.1	C2'-endo	C2'-endo	C2'-endo
C H A I N b	5'-C	143.4	C1'-exo	166.2	C2'-endo	150.3	C2'-endo	C2'-endo	C2'-endo
	G	47.9	C4'-exo	71.3	C4'-exo	47.8	C4'-exo	C4'-exo	C3'-endo
	C	150.8	C2'-endo	158.7	C2'-endo	157.1	C2'-endo	C2'-endo	C2'-endo
	G	80.8	O4'-endo	39.6	C4'-exo	40.3	C4'-exo	C3'-endo	C3'-endo
	C	149.5	C2'-endo	156.3	C2'-endo	150.9	C2'-endo	C2'-endo	C2'-endo
	G-3'	162.9	C2'-endo	167.8	C2'-endo	160.8	C2'-endo	C2'-endo	C2'-endo

5.3.2.2 Coordination of the Cu²⁺ ions

In **8OEA**, a total of eleven copper (II) ions were found after the native crystal was soaked in copper (II) chloride. The Cu²⁺ ions were added based on the electron density present and the distance from the N7 position of the guanine. All Cu²⁺ ions were coordinated to the N7 position of guanines. However, G_d6 was the only guanine in the structure with no coordinated copper (II) ion. After occupancy refinement, a mix of partial and total occupancies were attributed to all the Cu²⁺ ions as well as partial coordination spheres. Each Cu²⁺ ion was named following this convention: Cu²⁺ (G_{Chain}'Position') for clarity.

After soaking the crystal in a drop containing hydrogen peroxide, only three copper (II) ions were found to remain in **8OEB**. Likewise, the Cu²⁺ ions were added based on the electron

density present and the distance from the N7 position of the guanine. All Cu^{2+} ions were coordinated to the N7 position of guanines. After occupancy refinement, a mix of partial and total occupancies were attributed to all the Cu^{2+} ions as well as partial coordination spheres.

Finally, after a second soak of the same crystal in a drop containing hydrogen peroxide, a single copper (II) ion was identified in **8OEC**. Due to the low resolution of the data, a copper (II) ion was placed at this position due to the electron density, the proximity to the N7 position of the guanine and the low distance between the atom and a nearby water molecule. Details of copper coordination for all three structures are available in Table 5.7.

Surprisingly, the copper (II) ions disappeared from the crystal lattice as the experiment went on, only leaving a single copper (II) ion at low occupancy in **8OEC** (Figure 5.5.).

Table 5.7: Copper (II) ions parameters from 8OEA, 8OEB and 8OEC (distance in Å)

PDB ID	Metal name	Occupancy metal	Distance to guanine N7	Metal name	Occupancy metal	Distance to guanine N7
8OEA	Cu^{2+} (G _a 6)	0.93	2.14	Cu^{2+} (G _c 6)	0.72	2.14
	Cu^{2+} (G _a 4)	0.84	1.93	Cu^{2+} (G _c 4)	0.82	2.06
	Cu^{2+} (G _a 2)	1.00	2.19	Cu^{2+} (G _c 2)	0.51	2.08
	Cu^{2+} (G _b 6)	0.65	2.39	Cu^{2+} (G _d 4)	0.52	3.06
	Cu^{2+} (G _b 4)	0.58	2.13	Cu^{2+} (G _d 2)	0.92	2.54
	Cu^{2+} (G _b 2)	0.63	3.14			
8OEB	Cu^{2+} (G _a 6)	0.39	2.43			
	Cu^{2+} (G _a 4)	0.52	2.32			
	Cu^{2+} (G _a 2)	0.71	2.33			
8OEC	Cu^{2+} (G _a 2)	0.15	2.63			

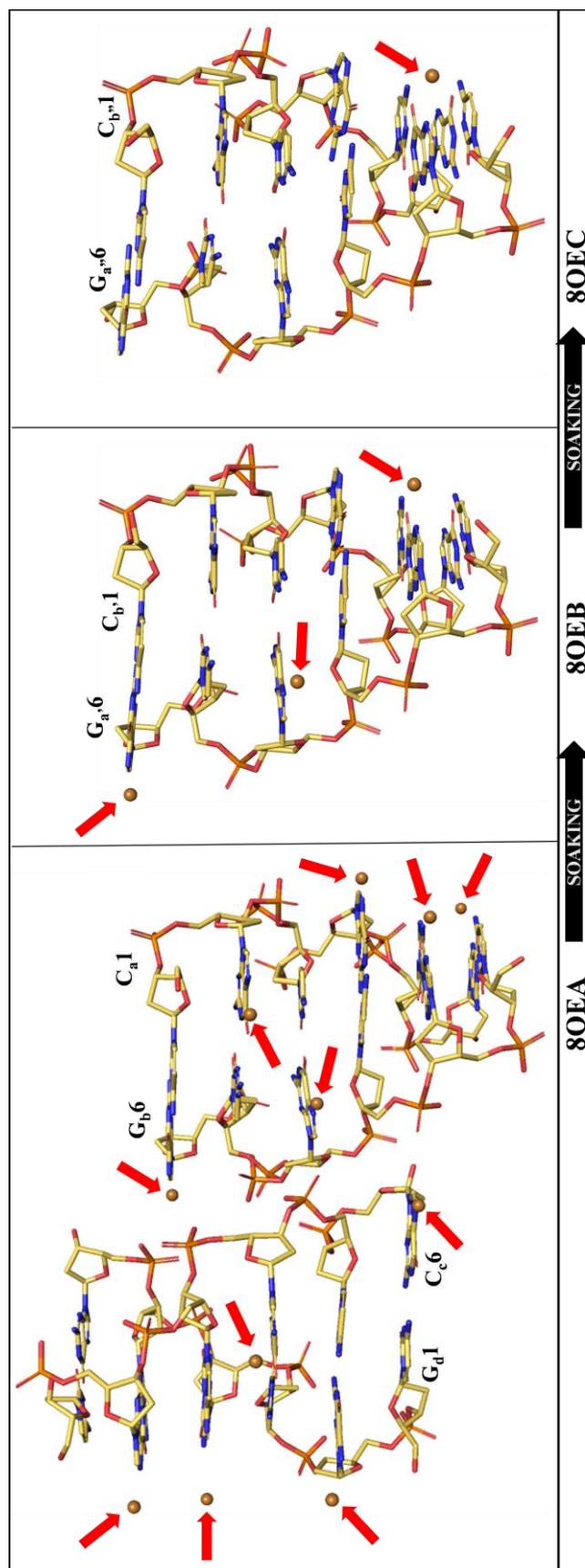


Figure 5.5: Representation of the disappearance of copper ions after soaking the same copper-DNA crystal with hydrogen peroxide twice. Eleven copper ions can be found in 80EA, then three in 80EB and only one in 80EC. Carbon, oxygen, nitrogen, and phosphate atoms are shown as yellow, red, blue, and orange respectively. Copper (II) ions are represented as orange spheres and are indicated by red arrows .

5.4. Discussion

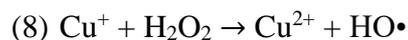
8OE7, **8OE8**, **8OE9**, **8OEB** and **8OEC** are all novel structures of the d(CGCGCG) sequence after a crystal was soaked in hydrogen peroxide. Although the original intent was to observe oxidative DNA damage in the crystal lattice with the hypothesis that an extra oxygen would appear at the C8 position of guanines, no evidence that suggests this type of damage occurred was observed in the electron density. The reaction between hydrogen peroxide and copper was expected to occur at the binding sites identified prior to hydrogen peroxide soaking and the hydroxyl radicals generated were expected to react with guanines to produce 8-oxoguanine, the major oxidation product during oxidative stress²³. However, no indication that any of the guanine bases were oxidised was reported. This could be due to the low conversion rate of this reaction in the absence of a reductant, as a conversion rate of 9.5 % was reported in these conditions¹⁴. Whilst this is potentially one explanation as to why damage was not observed, as crystal periodicity is a crucial characteristic to generate a diffraction pattern and hence, a 9.5 % conversion rate would likely be observable due to the high diffracting resolution of the crystals.

Whilst damage sites could not be identified, hydrogen peroxide molecules were found interacting with copper (II) ions in **8OE7** and **8OE8**, which is the first time this type of interaction has been observed in a DNA system. The average distance between a hydrogen peroxide molecule and copper was 1.85 Å. Previously, protein structures with at least a peroxide and a copper (II) ion in their active sites have been reported in the literature^{13,24–26} and distances between the two have been described. Some of the distances reported in **8OE7** and **8OE8** could be considered quite short compared to the average 1.90 - 2.00 Å which is typically observed^{13,25,27}.

Another disparity between the structures was the increase in the apparent disorder in the solvent channel. In **8OE7** and compared to what can be seen in **8OEY** at equivalent resolution, the electron density is not as well defined and might suggest a highly disordered solvent after peroxide soaking. Furthermore, the B-factors for the copper (II) ions ($\sim 60.0 \text{ \AA}^2$) were reported as six times higher than the average B-factors in the structure ($\sim 10.0 \text{ \AA}^2$), while the B-factors

for the hydrogen peroxide were not as high ($\sim 25.0 \text{ \AA}^2$). In comparison, the average B-factors for the copper (II) ions in **8OEY** was $\sim 10.0 \text{ \AA}^2$. B-factors in biological crystal structures at equivalent resolution are normally not as high. The B-factors for the copper (II) ions in **8OE7** are significantly above the maximal B-factor average values that are compatible with the crystalline solid state (B_{max}), described as 25 \AA^2 for resolution better than 1.5 \AA ²⁸. Large B-factors can be interpreted as a sign of internal motion, but also as a sign of crystal defect and diffraction decay²⁹. However, high B-factors were only reported in the solvent channel in this structure, and this could be explained by the ongoing reaction between the copper (II) ions and the hydrogen peroxide happening in the crystal, especially as the average B-factor in the structure was not described as high.

8OEA, **8OEB** and **8OEC** did not have ordered hydrogen peroxide molecules. However, it is worth noting how copper (II) ions are removed from the structure as the crystal is soaked in hydrogen peroxide. On average, there are five Cu (II) ions per duplex in **8OEA**, two and a half per duplex in **8OEB** and one per duplex in **8OEC**. As the reaction between the copper (II) ions and hydrogen peroxide goes on, the copper (II) ions are removed from their original binding sites. There could be several reasons why the copper (II) ions unbind from the guanines. Copper (II) ions are accessible to the solvent channel, and as indicated by the presence of H_2O_2 in **8OE7** and **8OES**, copper (II) chloride is reacting with the hydrogen peroxide present in the solvent channels. Copper (II) reacts with H_2O_2 to produce $\text{HO}\cdot$ (Scheme 5.2)³⁰. However, this reaction is not direct. Copper (II) first reacts with H_2O_2 to produce superoxide ions (O_2^-) and copper (I) ions (Scheme 5.2.(1)). There is a possibility that after copper (II) reacts with H_2O_2 , the copper (I) unbinds, then regenerates to copper (II) and binds again. However, Cu (I) was shown to also bind at the N7 position of guanine¹⁴. Yet, these two statements are not mutually exclusive, and it is possible that the structures described in this work are only intermediate of what is happening, and both copper (I) and copper (II) bind and unbind as the reaction goes on. Furthermore, the high B-factors could be an indication a lot of movement is happening while the reaction is ongoing.



Scheme 5.2: The reaction of copper with hydrogen peroxide

The complex mechanism presented here could explain the loss of copper (II) binding sites at the N7 position of the guanines and highlight why a disparity is observed between DNA oxidative damage caused by copper-catalysed Fenton and iron-catalysed Fenton chemistry. With constant binding and unbinding, the reaction centre in a copper-catalysed Fenton chemistry might not be as well defined as an iron-catalysed Fenton chemistry reaction centre, where iron (II) would directly react with hydrogen peroxide to produce hydroxyl radicals³. To further investigate this mechanism, crystals with ordered iron (II) centres could be soaked in hydrogen peroxide and compared to this system. This could help identify the differences between how iron (II) and copper (II) react with hydrogen peroxide in a DNA system.

5.5. References

1. Thenard, L. J. Observations sur des nouvelles combinaisons entre l'oxygène et divers acides. in *Annales de chimie et de physique* vol. 8 306–312 (1818).
2. Haber, F., Weiss, J. & Pope, W. J. The catalytic decomposition of hydrogen peroxide by iron salts. *Proc. R. Soc. Lond. Ser. - Math. Phys. Sci.* **147**, 332–351 (1997).
3. Fenton, H. J. H. LXXIII.—Oxidation of tartaric acid in presence of iron. *J. Chem. Soc. Trans.* **65**, 899–910 (1894).
4. Zuo, L., Zhou, T., Pannell, B. K., Ziegler, A. C. & Best, T. M. Biological and physiological role of reactive oxygen species – the good, the bad and the ugly. *Acta Physiol.* **214**, 329–348 (2015).
5. Brieger, K., Schiavone, S., Jr, F. J. M. & Krause, K.-H. Reactive oxygen species: from health to disease. *Swiss Med. Wkly.* **142**, 13659–13659 (2012).
6. Cooke, M. S., Evans, M. D., Dizdaroglu, M. & Lunec, J. Oxidative DNA damage: mechanisms, mutation, and disease. *FASEB J.* **17**, 1195–1214 (2003).
7. Hofer, T. & Perry, G. Nucleic acid oxidative damage in Alzheimer's disease—explained by the hepcidin-ferroportin neuronal iron overload hypothesis? *J. Trace Elem. Med. Biol.* **38**, 1–9 (2016).
8. Henle, E. S. *et al.* Sequence-specific DNA Cleavage by Fe²⁺-mediated Fenton Reactions Has Possible Biological Implications. *J. Biol. Chem.* **274**, 962–971 (1999).
9. Chevion, M. A site-specific mechanism for free radical induced biological damage: The essential role of redox-active transition metals. *Free Radic. Biol. Med.* **5**, 27–37 (1988).
10. Meneghini, R. Iron Homeostasis, Oxidative Stress, and DNA Damage. *Free Radic. Biol. Med.* **23**, 783–792 (1997).

11. Dillard, B. D., Demick, J. M., Adams, M. W. W. & Lanzilotta, W. N. A cryo-crystallographic time course for peroxide reduction by rubrerythrin from *Pyrococcus furiosus*. *JBIC J. Biol. Inorg. Chem.* **16**, 949–959 (2011).
12. Marangon, J. *et al.* Kinetic and Structural Studies of Aldehyde Oxidoreductase from *Desulfovibrio gigas* Reveal a Dithiolene-Based Chemistry for Enzyme Activation and Inhibition by H₂O₂. *PLOS ONE* **8**, e83234 (2013).
13. Shin, D. S. *et al.* Superoxide Dismutase from the Eukaryotic Thermophile *Alvinella pompejana*: Structures, Stability, Mechanism, and Insights into Amyotrophic Lateral Sclerosis. *J. Mol. Biol.* **385**, 1534–1555 (2009).
14. Fleming, A. M., Muller, J. G., Ji, I. & Burrows, C. J. Characterization of 2'-deoxyguanosine oxidation products observed in the Fenton-like system Cu(II)/H₂O₂/reductant in nucleoside and oligodeoxynucleotide contexts. *Org. Biomol. Chem.* **9**, 3338–3348 (2011).
15. Drozdal, P., Gilski, M., Kierzek, R., Lomozik, L. & Jaskolski, M. High-resolution crystal structure of Z-DNA in complex with Cr³⁺ cations. *JBIC J. Biol. Inorg. Chem.* **20**, 595–602 (2015).
16. Winter, G. xia2: an expert system for macromolecular crystallography data reduction. *J. Appl. Crystallogr.* **43**, 186–190 (2010).
17. McCoy, A. J. *et al.* Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674 (2007).
18. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2126–2132 (2004).
19. Afonine, P. V. *et al.* Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr. D Biol. Crystallogr.* **68**, 352–367 (2012).
20. Schrödinger, LLC. The PyMOL Molecular Graphics System, Version 1.8. (2015).

21. Zheng, G., Lu, X.-J. & Olson, W. K. Web 3DNA—a web server for the analysis, reconstruction, and visualization of three-dimensional nucleic-acid structures. *Nucleic Acids Res.* **37**, 240–246 (2009).
22. Macrae, C. F. *et al.* Mercury 4.0: from visualization to analysis, design and prediction. *J. Appl. Crystallogr.* **53**, 226–235 (2020).
23. White, B., Smyth, M. R., Stuart, J. D. & Rusling, J. F. Oscillating Formation of 8-Oxoguanine during DNA Oxidation. *J. Am. Chem. Soc.* **125**, 6604–6605 (2003).
24. Fukuda, Y. *et al.* Crystallographic study of dioxygen chemistry in a copper-containing nitrite reductase from *Geobacillus thermodenitrificans*. *Acta Crystallogr. Sect. Struct. Biol.* **74**, 769–777 (2018).
25. Rudzka, K. *et al.* Coordination of peroxide to the CuM center of peptidylglycine α -hydroxylating monooxygenase (PHM): structural and computational study. *JBIC J. Biol. Inorg. Chem.* **18**, 223–232 (2013).
26. Gabdulkhakov, A. *et al.* Investigations of Accessibility of T2/T3 Copper Center of Two-Domain Laccase from *Streptomyces griseoflavus* Ac-993. *Int. J. Mol. Sci.* **20**, 3184 (2019).
27. Jeoung, J.-H. *et al.* Bimetallic Mn, Fe, Co, and Ni Sites in a Four-Helix Bundle Protein: Metal Binding, Structure, and Peroxide Activation. *Inorg. Chem.* **60**, 17498–17508 (2021).
28. Carugo, O. How large B-factors can be in protein crystal structures. *BMC Bioinformatics* **19**, 61 (2018).
29. Sun, Z., Liu, Q., Qu, G., Feng, Y. & Reetz, M. T. Utility of B-Factors in Protein Science: Interpreting Rigidity, Flexibility, and Internal Motion and Engineering Thermostability. *Chem. Rev.* **119**, 1626–1665 (2019).
30. Halliwell, B., Adhikary, A., Dingfelder, M. & Dizdaroglu, M. Hydroxyl radical is a significant player in oxidative DNA damage *in vivo*. *Chem. Soc. Rev.* **50**, 8355–8360 (2021).

Chapter 6 - Conformational Changes Mediated by a Fenton-like Copper Reaction in a G-Quadruplex Forming Sequence.

6.1. Introduction

During normal cell metabolism, the regulation of reactive oxygen species (ROS) is important to keep levels appropriate to cell function. However, during oxidative stress, when ROS production increases and the cell's ability to neutralise ROS is diminished, many cellular components can be damaged and various disorders, including neurological disorders¹, cancer², and cardiovascular diseases³ can arise. Oxidative DNA damage has been reported as frequent in oxidative stress environment, with single strand breaks, double-strand breaks, the formation of apurinic/apyrimidinic lesions and base modifications among the type of damage described⁴. Often, oxidative DNA damage can lead to mutations and diseases⁴.

Out of all four DNA bases, guanine has the lowest redox potential⁵. As such, increased levels of oxidative damage have been reflected by the presence of the oxidative stress marker product 8-oxo-7,8-dihydroguanine (oxoG), a marker reported in the brains of individuals having suffered from neurodegenerative diseases and linked to metal dysregulation like Fe or Cu, transition metals known to react with hydrogen peroxide to generate hydroxyl radicals (HO[•])⁶. Previously, the role of guanine oxidation was not seen as anything else but the cause of mutations, however in recent years, research groups have started investigating the epigenetic role of oxidative damage and guanine oxidation⁷.

Found in the telomeric and gene promoter regions of the genome, G-quadruplexes form in G-rich nucleic acid sequences and consist of four guanine tetrads stabilised by Hoogsteen base-pairing interactions, stacked on top of each other and stabilised by the presence of sodium and/or potassium cations in the channel⁸ (Figure 6.1a). G-quadruplexes can form many different topologies and a multitude of proteins, including transcription factors, have been shown to recognise and bind specifically to quadruplex sequences^{9,10}. G-quadruplexes are also

particularly sensitive to oxidative damage^{11,12}. As a result, the presence of 8-oxoG in the promoter regions of genes has been investigated and subsequently, the suggestion that oxoG adopt a regulatory role during gene transcription via the base excision repair (BER) pathway arose¹³⁻¹⁶.

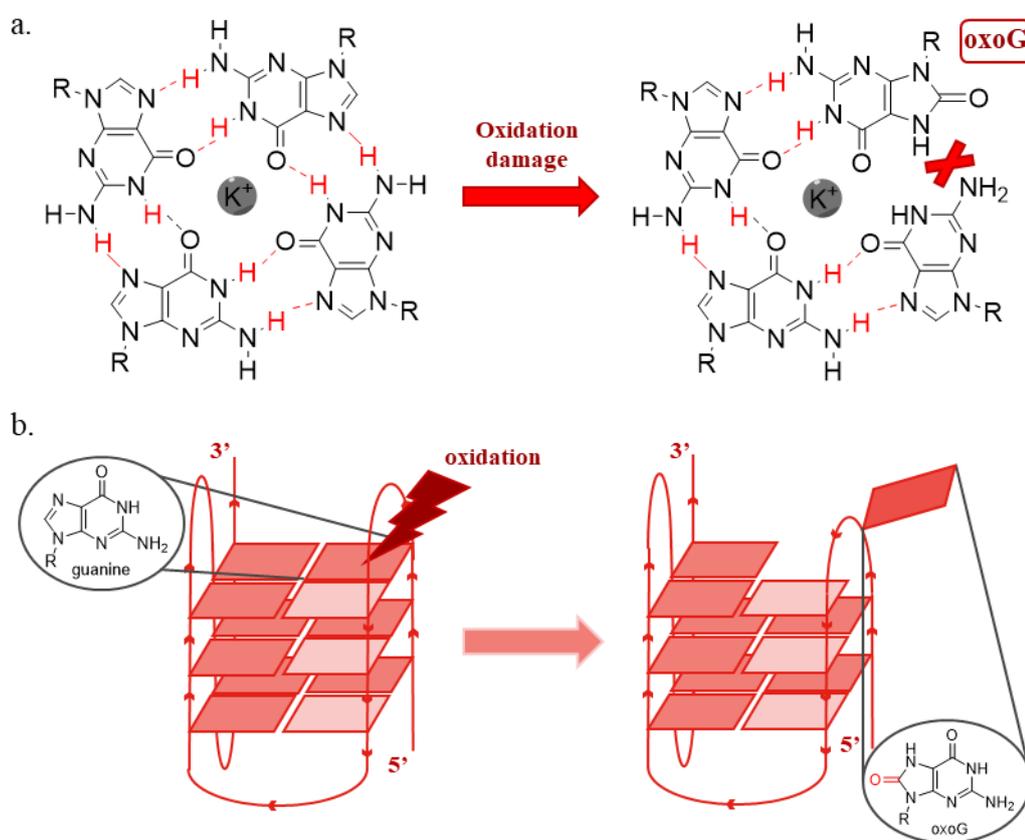


Figure 6.1: a. Schematic showing how a G-quadruplex is stabilized by Hoogsteen base-pairing with a potassium cation in the channel. After oxidative damage, pairing between the damaged guanine and the other guanine is prevented. b. The loss of hydrogen acceptor on the Hoogsteen edge of the damaged base can lead to structural destabilisation.

Various studies have looked at the effects of oxoG on G-quadruplexes structure¹⁷⁻²². And although this has provided some insight on what the effects structure are after damage, a single damaged base was incorporated prior to structural analysis in each study and the structural effects of damage in an actual oxidative stress environment were not examined.

Circular dichroism (CD) spectroscopy is a structural technique which can be used to differentiate the structure of biological molecules including DNA, in solution²³. Nucleic acids are chiral molecules which can absorb right- and left-handed circularly polarized light differently. This difference, commonly called circular dichroism, can be described with the quantity called ellipticity, Θ , expressed in degrees²³. Structural differences in DNA lead to unique CD spectral signatures, and with G-quadruplexes, specific CD spectral signatures have been associated with distinct G-quadruplex topologies: a positive band at 260 nm and a negative band at 245 nm for parallel G-quadruplexes, a positive band at 295 nm and a negative band at 260 nm for antiparallel quadruplexes, and two positive bands at 260 and 295 nm and a negative band at 245 nm for “hybrid” quadruplexes²⁴ (Figure 6.2a). These differences in topologies, arising from differences in arrangements of *anti/syn* glycosidic angles (Figure 6.2b), can be reliably tracked using CD and this technique can be used to monitor changes in the topology, especially when the guanine-tetrad stacks are affected as they mainly contribute to the CD signal²⁵.

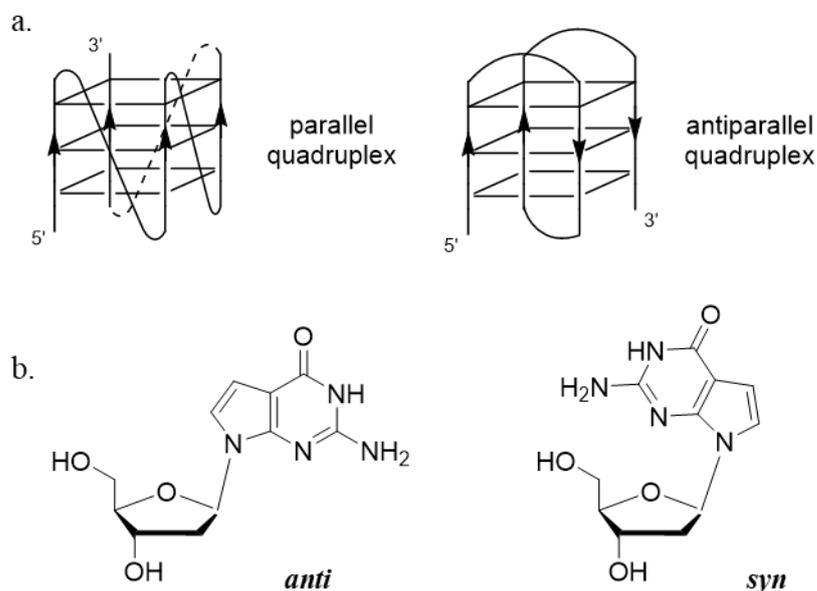


Figure 6.2: a. An example of possible parallel and antiparallel G-quadruplexes that have been reported. The name refers to the orientation of the strands part of the quadruplex core. b. Representation of the *anti* and *syn* glycosidic bond angles that can be adopted by guanine bases in G-quadruplexes.

Hence, changes in G-quadruplex topologies during oxidative damage can be monitored, as one of the consequences of guanine oxidation is the protonation at the N7 position of guanines, a modification which can prevent Hoogsteen pairings between the oxidised guanine and the other guanines in the tetrad and destabilise the whole structure^{19,26} (Figure 6.1b).

In this work, circular dichroism (CD) was used to monitor changes in a G-quadruplex structure as a function of time and temperature with the expectation that oxidative damage to guanines will prevent the guanine tetrads from reforming. For this, a G-quadruplex forming sequence found in the P1 promoter region of the PSEN2 gene, a gene which encodes for the presenilin-2 protein known for its role in processing amyloid precursor protein²⁷, was chosen. PSEN2 expression is mainly controlled by the P1 promoter whose activity depends on two transcription binding sites for the transcription factor Sp1²⁸ (Figure 6.3).

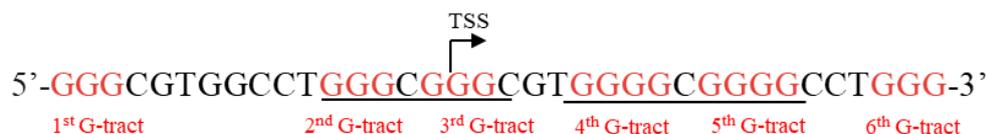


Figure 6.3: Transcription start site (TSS) of the PSEN2 gene. This exact sequence was used in this study. The binding sequences of SP1 have been underlined. Each track of guanines hypothesised to be part of the core is represented in red.

Considering past studies have established how OxoG in GC-rich Sp1 binding sites enhance gene transcription¹⁵, this sequence appeared as an excellent candidate to monitor oxidative damage. The 36-nucleotide long sequence was chosen after the gene promoter sequence was analysed using QGRS Mapper²⁹ to predict which portion of the promoter was more likely to form a G-quadruplex. After the sequence was prepared in a potassium cacodylate buffer, and the quadruplex formation was validated with CD, the sequence was confirmed to fold into a hybrid G-quadruplex in presence of potassium (Figure 6.4).

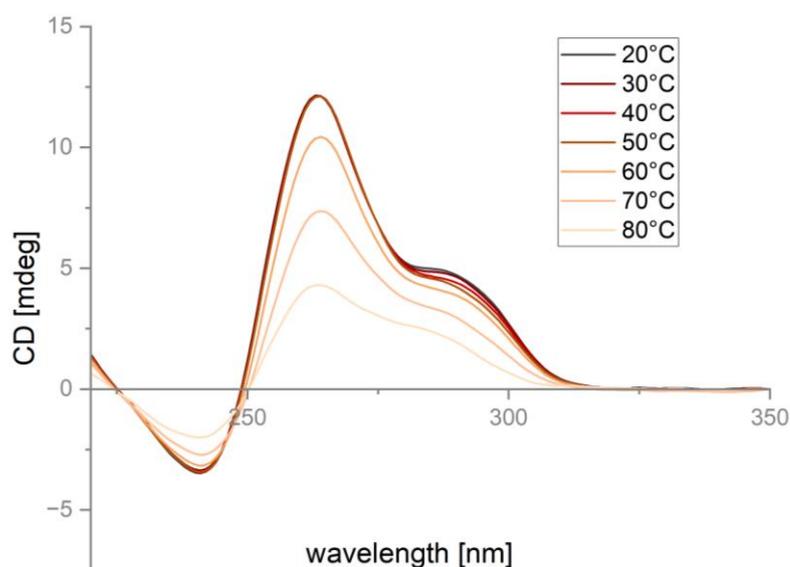


Figure 6.4: CD spectra of PSEN2 in 20mM potassium cacodylate buffer at various temperature. The PSEN2 sequence folds into a hybrid G-quadruplex in presence of potassium.

Separately, sequences of PSEN2 which had their tract of guanines (G-tracts) substituted by either A-tract or T-tract (Table 6.1) were investigated to identify which tracts participate in the tetrad core as it is unclear which G-tracts are implicated and only four out of the six found in the PSEN2 sequence are needed to form the G-quadruplex. Ultimately, this should also help identify which G-tracts could potentially be damaged primarily in an oxidative stress environment, as adenine and thymine bases will not participate in Hoogsteen base-pairing in the tetrad core, the same way oxidised guanines will not be able to form Hoogsteen base pairs with other guanines.

Table 6.1: The sequences used in this study. A is adenine, C is cytosine, G is guanine and T is thymine. Each tract of guanines (G-tract) hypothesised to be part of the G-quadruplex core is represented in red. The G-tracts which have been substituted by either A-tracts or T-tracts are represented in blue.

Name	Sequence (5' to 3')
PSEN2	GGGCGTGGCCTGGGCGGGCGTGGGGCGGGGCCTGGG
G3T4	GGGTTTGGGTTTTGGGTTTTGGGTTTT
PSEN2 A1	AAACGTGGCCTGGGCGGGCGTGGGGCGGGGCCTGGG
PSEN2 A2	GGGCGTGGCCTAAACGGGCGTGGGGCGGGGCCTGGG
PSEN2 A3	GGGCGTGGCCTGGGCAAAACGTGGGGCGGGGCCTGGG
PSEN2 A4	GGGCGTGGCCTGGGCGGGCGTAAAAAGGGGCCTGGG
PSEN2 A5	GGGCGTGGCCTGGGCGGGCGTGGGGCAAAACCTGGG
PSEN2 A6	GGGCGTGGCCTGGGCGGGCGTGGGGCGGGGCCTAAA
PSEN2 T1	TTTCGTGGCCTGGGCGGGCGTGGGGCGGGGCCTGGG
PSEN2 T2	GGGCGTGGCCTTTTCGGGCGTGGGGCGGGGCCTGGG
PSEN2 T3	GGGCGTGGCCTGGGC TTTCGTGGGGCGGGGCCTGGG
PSEN2 T4	GGGCGTGGCCTGGGCGGGCGT TTTTCGGGGCCTGGG
PSEN2 T5	GGGCGTGGCCTGGGCGGGCGTGGGGC TTTTCCTGGG
PSEN2 T6	GGGCGTGGCCTGGGCGGGCGTGGGGCGGGGCCTTTT

Subsequently, Fenton reagents (H₂O₂ 1 mM, CuCl₂ 200 μM, citric acid 1 mM) were added to samples containing the DNA sequence PSEN2, which were monitored using CD, to look at the effects of oxidative damage. The DNA sequence G3T4 (Table 6.1), a parallel G-quadruplex forming sequence containing only four G-tracts, was also used in this experiment to monitor the effects of oxidative damage on a model system. Finally, a sample of the DNA sequence PSEN2 treated with the Fenton reagents to induce oxidative damage was analysed on a denaturing gel to identify potential strand breaks.

This method was developed in parallel with the method that was presented in Chapter 5 and the rest of the work presented in this thesis. The overall aim was to find interesting systems with the method presented here and to attempt crystallisation to further investigate mechanism of oxidative damage as a function of structure.

6.2. Material and methods

6.2.1. DNA and samples preparation

The oligonucleotides were purchased from Eurogentec as a solid purified by RP-HPLC. DNA concentration was determined from the value of the absorbance at $\lambda = 260$ nm, using the molar extinction coefficient provided by the manufacturer. All sequences were prepared to 200 μ M stock solutions with HPLC water. Stock solutions of the chemicals used for the Fenton experiment were freshly prepared before each experiment: potassium chloride (1 M), copper (II) chloride (50 mM), citric acid (200 mM), potassium cacodylate buffer (pH 7.0, 200 mM), potassium sulfite (250 mM) and hydrogen peroxide (1% w/v) (SLS). DNA samples were annealed in a 20 mM potassium cacodylate pH 7.0 buffer and 50 mM potassium chloride at 358K for 2 minutes then slowly cooled down at RT. All chemicals were purchased from Sigma-Aldrich unless stated otherwise.

6.2.2. Quadruplexes formation and melting experiments.

Circular dichroism spectra were obtained using a Chirascan V100 – Circular dichroism spectrometer. Spectra were collected from 200-350 nm with a 1 second integration time, 1 nm slit and a reduced volume quartz cell with a 1 mm path length. Data were collected every 10 °C Celsius from 20 to 80 degrees Celsius at a ramp-rate of 2 degrees Celsius per minute then cooled down to 20 degrees Celsius. All spectra of the unmodified PSEN2 sequence were obtained using a DNA concentration of 15 μ M. All spectra of substituted PSEN2 sequences (adenine or thymine) were obtained at a DNA concentration of 20 μ M. All spectra were corrected by subtracting the background and by correcting the offset at 350 nm. Table 6.1 lists all the sequences used in this study.

6.2.3. Fenton reaction monitoring using Circular Dichroism

Circular dichroism spectra were obtained using a Chirascan V100 – Circular dichroism spectrometer. Spectra were collected from 200-350 nm with a 1 second integration time, 1 nm slit and a reduced volume quartz cell with a 1 mm path length. Copper (II) chloride, citric acid and hydrogen peroxide were added to annealed samples to a concentration of 200 μM , 1 mM and 1 mM respectively. The hydrogen peroxide was added to the cuvette containing the sample last and data collection was started immediately after. Data were collected every 10 degrees Celsius from 20 to 80 degrees Celsius then from 80 to 20 degrees Celsius at a ramp-rate of 2 degrees Celsius per minute to monitor the Fenton reaction and DNA damage. All spectra were obtained for a constant DNA concentration (15 μM). All spectra were corrected by subtracting the background and by correcting the offset at 350 nm.

6.2.4. Sample preparation for analysis

Samples were quenched using potassium sulfite and desalted using ethanol precipitation. 3 volumes of ice cold 100 % ethanol were directly added to the samples, mixed thoroughly, and left at -20 °C overnight. The samples were then centrifuged at 13,000 rpm for 30 minutes at 4 °C. The supernatants were removed, and the pellets were washed carefully with ice cold 75 % ethanol and spun at 13,000 rpm for 10 minutes. This step was repeated one more time and the supernatants removed. The pellets were then left to air dry and resuspended in deionised water. Samples were made to ensure 200 ng are loaded on the gels. To ensure separation on the denaturing gel, three DNA sequences of various lengths were prepared alongside these samples to act as reference samples.

6.2.5. Preparation of denaturing polyacrylamide gels

To analyse the samples, an 18 % urea (8M) polyacrylamide gel was prepared. 4 mL of 10x TBE buffer, 18 mL of a 40 % acrylamide: bis-acrylamide 29:1 solution, 19.2 g of urea and HPLC water were mixed to make a 40 mL gel solution. The solution was shaken vigorously until the urea was dissolved. 40 μL of TEMED was added to the mix, followed by 400 μL of freshly prepared 10% (w/v) APS. The solution was mixed, poured between the gel plates and a comb was placed immediately before the gel starts polymerising. The gel was allowed to

polymerise for 30 minutes. Once complete, the comb was removed, and the reservoir of the electrophoresis tank was filled with 1x TBE buffer. Gel was pre-run at 125 V for 30 minutes, then the wells washed before loading the samples. The samples were mixed with 6x Bromophenol Blue and loaded on the gel. The electrophoresis tank was placed in an oven at 60 °C and run for one hour at 150 V. The gel was then soaked in 1x TBE for 15 minutes, followed by staining in a 1x TBE solution mixed with SYBR Gold. The gel was then examined under UV light.

6.3. Results

6.3.1. Confirmation of G-quadruplex formation by circular dichroism

CD measurements were first run to confirm the PSEN2 sequence forms a hybrid G-quadruplex in presence of potassium. Then, to ensure the G-quadruplex folds and unfolds correctly during the experiments, a protocol identical to the one that will be used to monitor samples treated with Fenton reagents was used. The temperature was slowly increased to 80 °C, then decreased back to 20 °C at the same rate. CD spectra were collected every 10 °C and the result can be seen in Figure 6.5. The CD signal at 240, 265 and 295 nm recovered after cooling the sample to 20 °C, which suggests the PSEN2 sequence folds back correctly into a “hybrid” G-quadruplex under the time constraint created by the experiment.

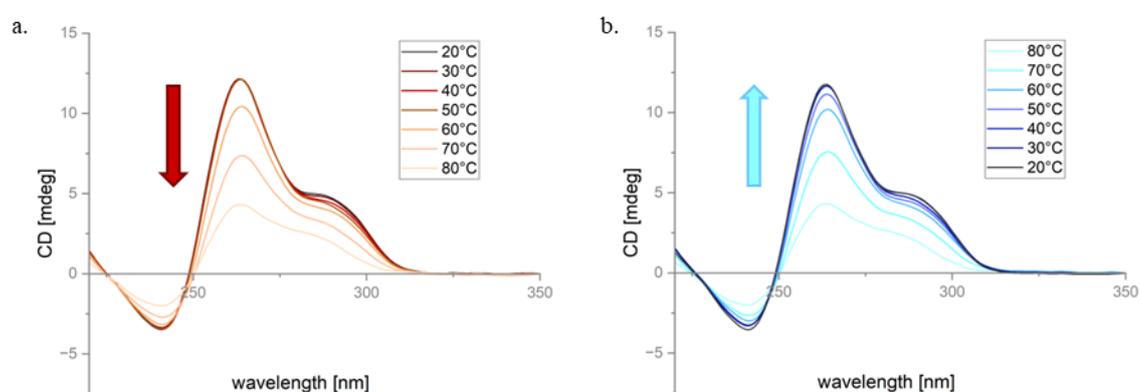


Figure 6.5: CD spectra of the PSEN2 G-quadruplex as a function of temperature. a. As the temperature increases, the G-quadruplex is denatured. b. Decreasing the temperature favoured the refolding of the G-quadruplex in its original “hybrid” topology.

6.3.2. Effects of oxidative damage on PSEN2 topology after thermal denaturation

Before the run, copper (II) chloride and citric acid were added to the annealed sample and transferred to a 1 mm cuvette. In this study, copper was used as a non-oxygen sensitive model for iron and citric acid was used as a substitute to ascorbic acid, a chiral molecule with a strong CD signal. To coincide the addition of hydrogen peroxide with the start of the experiment, the hydrogen peroxide was added to the cuvette last and data collection was started immediately after. At the end of the experiment, the sample was immediately quenched with potassium sulfite and desalted using ethanol precipitation.

While the Fenton reaction was ongoing, the temperature was increased and decreased at the same rate used to ensure the G-quadruplex could refold with the same topology. Consequently, the two results can be directly compared. The major consequence of denaturing the PSEN2 G-quadruplex in the presence of Fenton reagents was the changes in the CD spectra between the start (20 °C, native, Figure 6.6a) and the end (20 °C, damaged/undamaged, Figure 6.6a) of the experiment. A major 48.6 % decrease in the ellipticity at 265 nm was observed, a result that was anticipated as oxidative damage promotes the destabilisation of G-quadruplexes. A similar decrease in the ellipticity was observed when the same method was employed with the G3T4 sequence (Supplementary information). However, an unexpected 17.0 % increase in the ellipticity at 295 nm was noticed, which could suggest the sequence refolds differently, an increase in the signal at 290 nm could suggest a structure with a more prominent anti-parallel component (Figure 6.6a).

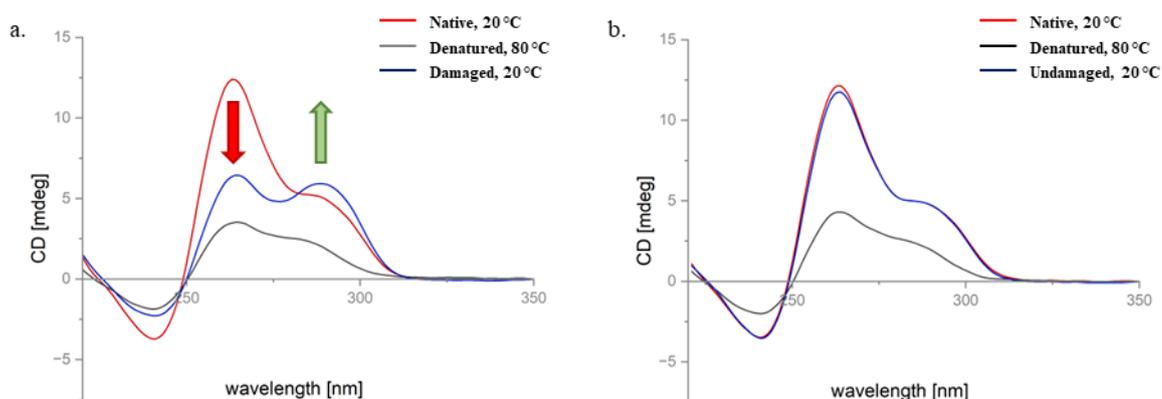


Figure 6.6: a. CD spectra of the PSEN2 G-quadruplex before (red) and after (blue) the Fenton reagents (H_2O_2 1 mM, CuCl_2 200 μM , citric acid 1 mM) were added to the sample. The sample was slowly brought to 80 °C, a temperature which would have denatured the DNA. After cooling down the sample back to 20 °C, changes in the CD spectra were identified. A decrease in the ellipticity at 265 nm and an increase in the ellipticity at 295 nm were observed, which could suggest the sequence refolds differently. b. In comparison, no changes in the CD spectra were identified when no Fenton reagents were added to the and the DNA was able to refold in its original “hybrid” topology at the end of the experiment.

6.3.3. Effects of G-tract substitution on PSEN2 topology

The PSEN2 sequence contains six G-tracts, but only four out of the six will participate in the tetrad core, while the other two will be part of the loop region. To identify which G-tracts are part of the core and possibly understand which G-tracts have potentially been damaged, CD measurements of twelve substituted sequences were run at 20 °C. Each sequence has had one of the six G-tracts substituted by either T-tracts or A-tracts. All the substituted PSEN2 sequences and given names are summarised in Table 6.1.

The native PSEN2 sequence forms a “hybrid” G-quadruplex in the presence of potassium, with distinct positive bands at 260 and 295 nm and a negative band at 240 nm (Figure 6.4). And when compared to the CD profile of the **PSEN2 A1**, **PSEN2 A6**, **PSEN2 T1**, and **PSEN2 T6** sequences, no significant differences were observed. The same positive bands at 265 and 295 nm and negative band at 240 nm (Figure 6.7) were identified, a result which suggests that substituting the first and sixth G-tracts did not prevent these DNA sequences from folding into the same hybrid G-quadruplexes than the native G-quadruplex, regardless of whether the substitution was done with A-tracts or T-tracts.

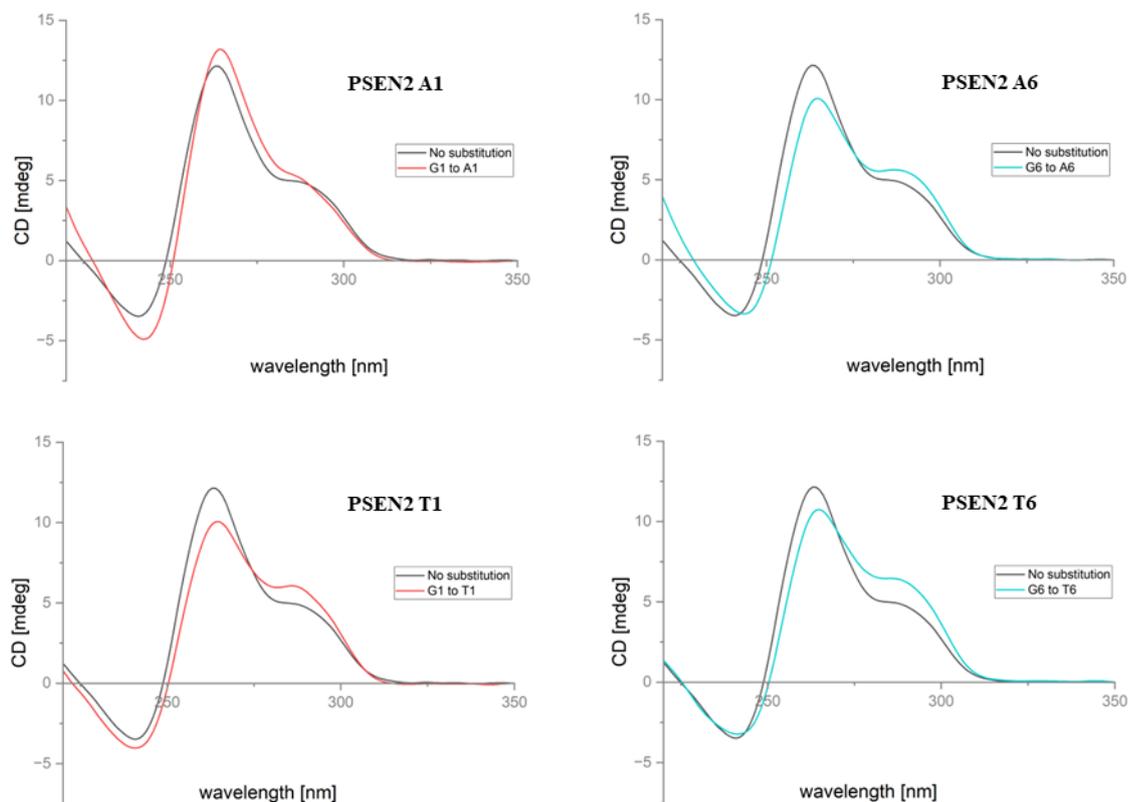


Figure 6.7: CD spectra of all the substitutions in the PSEN2 sequence resulting in no significant topological changes. 1st and 6th G-tract substitutions did not influence topology, regardless of if the substitution was done with adenine or thymine bases. Each sample was prepared using 20 mM potassium cacodylate pH 7.0 buffer and 50 mM potassium chloride. Each CD spectra was collected at 20 °C and was compared to the CD spectra of the native PSEN2 sequence.

Additionally, substitution of the third G-tract resulted in substantial variations in the band intensity at 295 nm, while the same positive band at 265 nm and negative band at 240 nm found in the CD spectra of the native sequence were observed, suggesting both sequences still formed a “hybrid” G-quadruplex. The significant increase in the band intensity at 295 nm in **PSEN2 A3** and **PSEN2 T3** can be seen in Figure 6.8. Similarly, substitution of the fourth G-tract resulted in positive bands at 260 and 295 nm and a negative band at 240 nm, suggesting the formation of a “hybrid” G-quadruplex as well, except the signal is much stronger at 295 nm than at 260 nm, which could suggest a shift towards a “hybrid” structure with a stronger anti-parallel signal. The CD profiles of the **PSEN2 A4** and **PSEN2 T4** sequences can be seen in Figure 6.8.

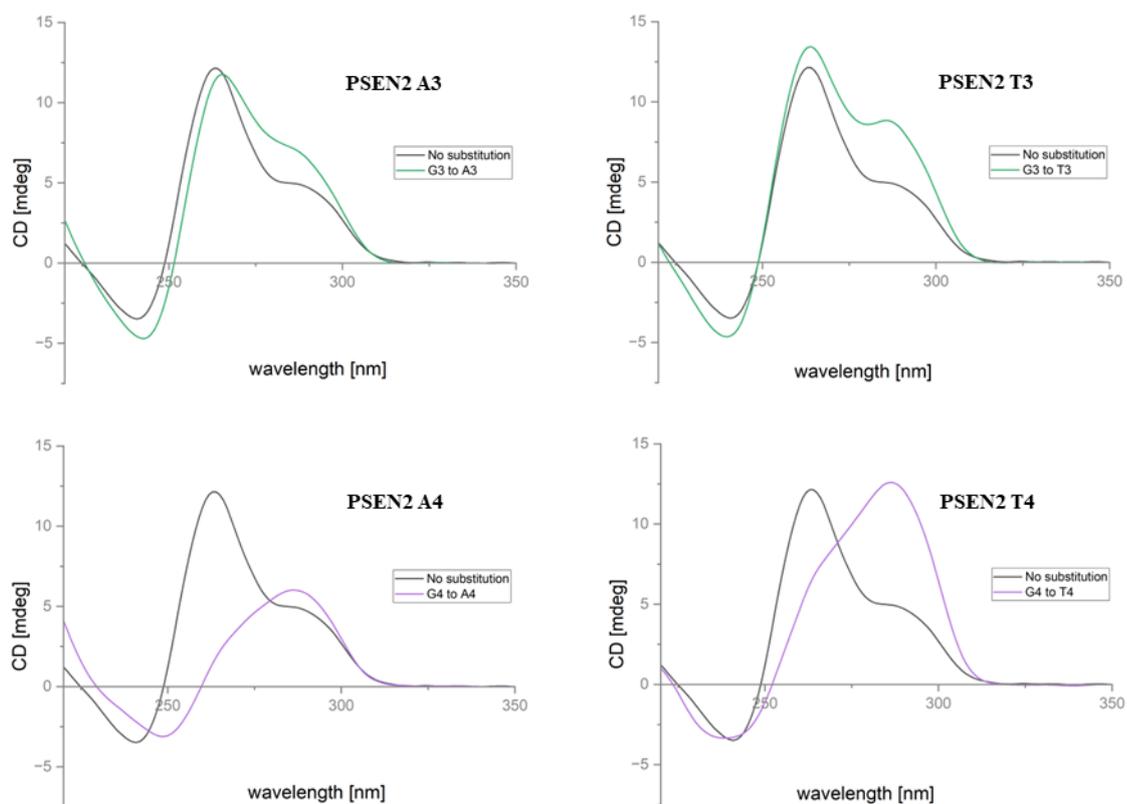


Figure 6.8: CD spectra of all the substitutions in the PSEN2 sequence resulting in small changes which did not change the overall topology. 3rd G-tract substitutions was not influenced by whether the substitution was done using adenine or thymine bases. 4th G-tract substitutions led to a stronger CD signal at 290 nm, but a positive band at 260 nm suggests this sequence still forms a “hybrid” quadruplex. Each sample was prepared using 20 mM potassium cacodylate pH 7.0 buffer and 50 mM potassium chloride. Each CD spectra was collected at 20 °C and was compared to the CD spectra of the native PSEN2 sequence.

Finally, major changes in band intensity were detected in **PSEN2 A2**, **PSEN2 A5**, **PSEN2 T2**, and **PSEN2 T5**. The CD spectra all displayed an intensive positive band at 295 nm and a negative one at 265 nm (Figure 6.9), consistent with the presence of an antiparallel G-quadruplex in solution.

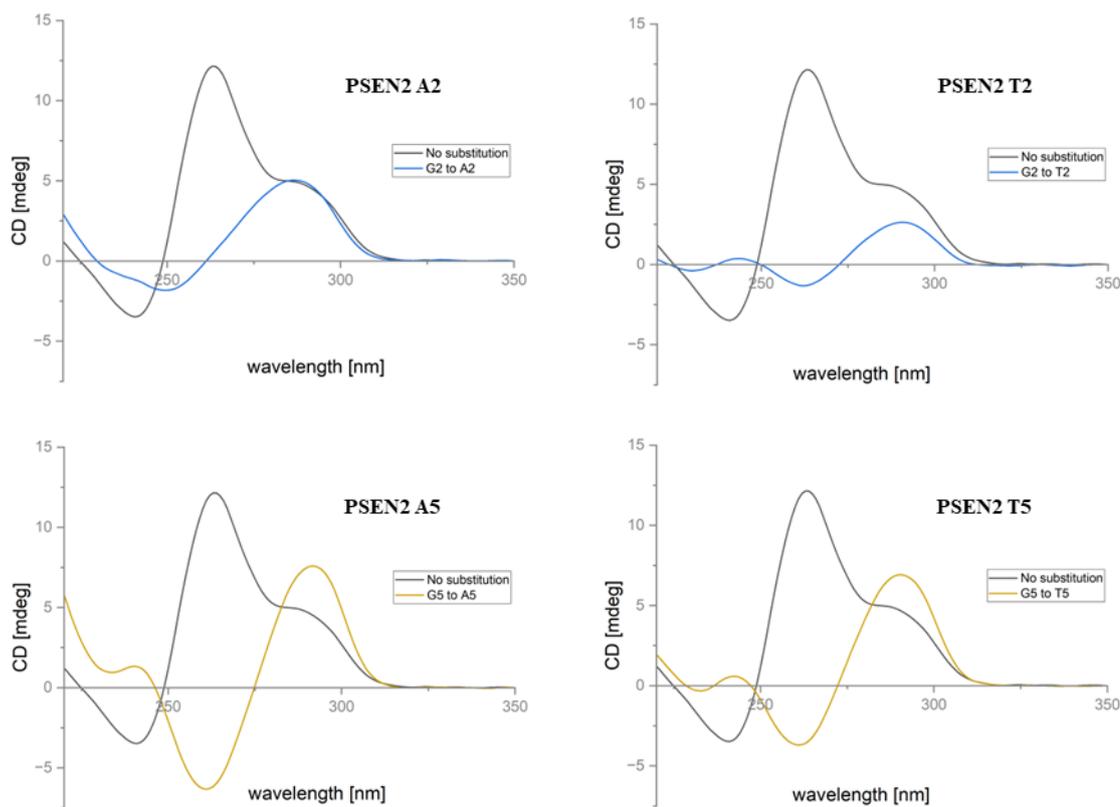


Figure 6.9: CD spectra of all the substitutions in the PSEN2 sequence resulting in the change of the overall topology. 2nd and 5th G-tract substitutions led to the formation of an anti-parallel G-quadruplex. The change was not influenced by the type of substitution (adenine or thymine). Each sample was prepared using 20 mM potassium cacodylate pH 7.0 buffer and 50 mM potassium chloride. Each CD spectra was collected at 20 °C and was compared to the CD spectra of the native PSEN2 sequence.

6.3.4. Identifying strand breaks after the Fenton reaction

Although over half of the damage done to DNA by hydroxyl radicals accounts to damage to bases³⁰, strand breaks still constitute a significant portion of the damage mediated by Fenton reaction^{31,32}. To ensure the conformational change observed after denaturing and renaturing the sample treated with Fenton reagents were not due to strand cleavage, the samples were run on a denaturing urea polyacrylamide gel. Denaturing polyacrylamide gels allow for the separation of DNA molecules based on their length³³; thus, strand scission could be identified by the presence of multiple resolved bands representing the smaller cleaved DNA strands. The PSEN2 sequence is 36-nucleotides long, so being able to observe band separation between relatively shorts strands was important. To this purpose, a high percentage gel (18 %) was used. And to ensure resolution was appropriate, a mix of three nucleotides of different lengths was loaded on the gel.

After the experiment, a first single band was identified as the untreated PSEN2 sample and was used as a control to determine how far a 36-nucleotide long sequence migrates in these conditions. A second single band was identified as the PSEN2 sample treated with Fenton reagents and no strand breaks were identified (Figure 6.10).

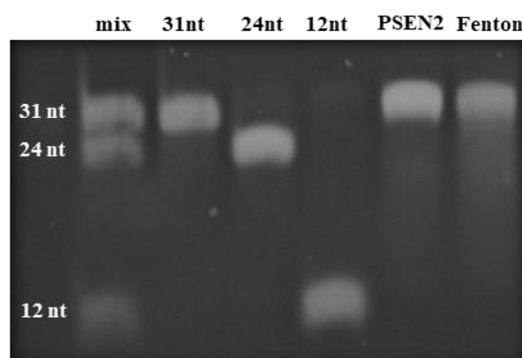


Figure 6.10: 18% denaturing urea polyacrylamide gel (urea, 8M). PSEN2 samples treated with Fenton reagents were desalted using ethanol precipitation. Gel was run at 150V for one hour. Four reference samples were used to good resolution: a 31-nucleotide long DNA sequence, a 24-nucleotide long DNA sequence and a 12-nucleotide long DNA sequence. A sample containing all three DNA references was loaded on the gel to ensure separation. The DNA samples were all suspended in deionised water before being loaded on the gel.

6.4. Discussion

In this study, a G-quadruplex forming sequence found in the promoter region of the PSEN2 gene was treated with Fenton reagents and conformational changes were monitored using Circular Dichroism. G-quadruplexes are particularly known to be susceptible to oxidative damage^{11,12} and many studies have looked at the structural effects of damage on G-quadruplexes. One of the consequences of guanine oxidation is the protonation at the N7 position of guanines, a modification which can prevent pairing between bases and knock damaged G-tracts out of the core tetrads^{19,26}. As a result, several mechanisms involving damaged G-tracts and conformational changes, like how oxoG can regulate gene expression when incorporated into G-quadruplex forming sequences found in promoter, have been reported³⁴⁻³⁶. More recently, the “spare tire” mechanism was described, where a fifth tract of guanines is recruited to aid in the repair of damaged G-quadruplexes¹⁶. These spare G-tracts

were portrayed as facilitating the repair process, to prevent gene dysregulation in the event guanines involved in the G-quadruplex's core formation are damaged under oxidative stress conditions. However, Brown *et al.* later concluded extra guanine-runs were less likely to act as “spare tires” but were more likely to be part of a complete system in which equilibrating G-quadruplex structures exist³⁷.

The epigenetic role of oxoG is still quite misunderstood, but an increasing number of studies about the effects of oxoG in G-quadruplexes, and more particularly in G-quadruplex forming sequences found in promoter regions, is being published^{18,26,38}. In this Chapter, the G-quadruplex forming sequence PSEN2 is being investigated. PSEN2 is a 36-nucleotides long sequence found in the promoter region of the PSEN2 gene. Previously, alteration in the PSEN2 gene's expression has been described as a risk factor for Alzheimer's disease (AD)³⁹. Furthermore, increased levels of oxidative damage have been reported in the brains of individuals having suffered from AD⁶. Therefore, the hypothesis that the presence of oxidative damage in the promoter region of the PSEN2 gene regulate gene expression can be explored.

The sequence was characterised with CD and was found to form a hybrid quadruplex in the presence of potassium cations. A CD melt was initially run to control that the sequence can unfold and refold correctly in the time the experiment takes, and it was determined that PSEN2 does not undergo conformational changes under these conditions. Then, the same experiment was run with PSEN2 mixed with Fenton reagents, and this time, a significant decrease in the signals at 245 and 260 nm was observed. As expected, oxidative damage can destabilise the whole structure and a loss of CD signal is expected as the CD signal originates from the stacking interactions between the guanines⁴⁰. However, an increase in the CD signal at 295 nm was not expected in this scenario, as the whole structure should have been destabilised. An increase in the signal at 295 nm in this case could suggest the reformation of predominantly anti-parallel G-quadruplexes. Multiple studies have described how the presence of oxoG has induced changes in the CD spectra^{16,18,26,38}. For example, changes in the CD spectra of VEGF were reported after oxoG was incorporated in the core and loop region of the quadruplex. These changes were attributed to the presence of a “spare tire” which stabilises the quadruplex formation¹⁶. In this work, the presence of two potential “spare tires” in the sequence could

explain why this sequence was able to refold, as these “spare tires” would aid in the repair of the G-quadruplex.

However, CD does not provide structural information at an atomic level, hence the localisation of the damaged bases cannot be identified with this method alone. To understand where oxidative damage could potentially happen, each G-tract was substituted by adenine and thymine bases, as they will not participate in Hoogsteen base-pairing in the tetrad core. These substitutions can help understand the structural consequences following the damage of entire tracts of guanines in an oxidative stress environment, as well as potentially identify which G-tracts are implicated in the tetrad core.

The CD spectra of all the substitutions were observed and a mixture of different conformations were described. When the first and sixth G-tracts were substituted in the PSEN2 sequence, no significant changes were noted in these CD spectra compared to the CD spectra of the native sequence. This could suggest that these guanine bases are not implicated in the tetrad core in the native structure. Other substitutions such as the substitutions of the third and fourth G-tracts by adenine and thymine bases led to changes in the CD spectra compared to the native CD spectra, especially in the signal at 295 nm. However, these sequences still form hybrid G-quadruplexes, albeit with a stronger anti-parallel signal. The last substitutions, second and fifth G-tracts, were found to have a stronger effect on the topology of the G-quadruplexes. The positive signal at 295 nm and negative signal at 260 nm observed in these CD spectra are indicative of the formation of anti-parallel quadruplexes.

None of the CD spectra observed after the substitutions resembled the CD spectra collected after treating the PSEN2 sequence with Fenton reagents. However, specific substitutions have induced shifts to different topologies and multiple G-quadruplex topologies could possibly exist in an equilibrium in solution³⁷. Although entire G-tracts were substituted in this case, and not single guanine-bases, oxidative damage could still lead to a similar equilibrium in solution. The diversity in conformations observed from the substituted sequences might also indicate oxidative damage can damage more than one tract of guanines and a mix of

different conformation is present in the sample. The result from the denaturing gel indicate strand break is unlikely, a result that supports the hypothesis that a mix of conformations might be present in the sample. These conformational changes, especially when the sequence is present in the promoter, could lead to the sequence no longer being recognised by the transcription factor SP1, whose binding sites is present in the PSEN2 sequence²⁸. The presence of oxoG in an SP1 binding site has been shown to enhance gene transcription in the past and it is possible a similar mechanism exists for PSEN2¹⁵. However, further experiments done with 8-oxoguanine substitutions would still be needed to confirm the effect of damaged guanines in the PSEN2 sequence, for a more adequate model. Mass spectrometry and High-performance liquid chromatography (HPLC) can also be used to determine the presence of oxidative damage-products post DNA-oxidation⁴¹⁻⁴³. These techniques could be applied to identify oxoG bases after treating the PSEN2 sequence with Fenton reagents.

6.5. References

1. Zuo, L., Zhou, T., Pannell, B. K., Ziegler, A. C. & Best, T. M. Biological and physiological role of reactive oxygen species – the good, the bad and the ugly. *Acta Physiol.* **214**, 329–348 (2015).
2. Liou, G.-Y. & Storz, P. Reactive oxygen species in cancer. *Free Radic. Res.* **44**, 479–496 (2010).
3. Brieger, K., Schiavone, S., Miller, F. J. & Krause, K.-H. Reactive oxygen species: from health to disease. *Swiss Med. Wkly.* **142**, w13659 (2012).
4. Cooke, M. S., Evans, M. D., Dizdaroglu, M. & Lunec, J. Oxidative DNA damage: mechanisms, mutation, and disease. *FASEB J. Off. Publ. Fed. Am. Soc. Exp. Biol.* **17**, 1195–1214 (2003).
5. Prat, F., Houk, K. N. & Foote, C. S. Effect of Guanine Stacking on the Oxidation of 8-Oxoguanine in B-DNA. *J. Am. Chem. Soc.* **120**, 845–846 (1998).
6. Hofer, T. & Perry, G. Nucleic acid oxidative damage in Alzheimer’s disease—explained by the hepcidin-ferroportin neuronal iron overload hypothesis? *J. Trace Elem. Med. Biol.* **38**, 1–9 (2016).
7. Giorgio, M., Dellino, G. I., Gambino, V., Roda, N. & Pelicci, P. G. On the epigenetic role of guanosine oxidation. *Redox Biol.* **29**, 101398 (2020).
8. Rhodes, D. & Lipps, H. J. G-quadruplexes and their regulatory roles in biology. *Nucleic Acids Res.* **43**, 8627–8637 (2015).
9. Kendrick, S. & Hurley, L. H. The role of G-quadruplex/i-motif secondary structures as cis-acting regulatory elements. *Pure Appl. Chem.* **82**, 1609–1621 (2010).
10. Huppert, J. L. & Balasubramanian, S. G-quadruplexes in promoters throughout the human genome. *Nucleic Acids Res.* **35**, 406–413 (2007).

11. Sugiyama, H. & Saito, I. Theoretical Studies of GG-Specific Photocleavage of DNA via Electron Transfer: Significant Lowering of Ionization Potential and 5'-Localization of HOMO of Stacked GG Bases in B-Form DNA. *J. Am. Chem. Soc.* **118**, 7063–7068 (1996).
12. Ohno, M. *et al.* A genome-wide distribution of 8-oxoguanine correlates with the preferred regions for recombination and single nucleotide polymorphism in the human genome. *Genome Res.* **16**, 567–575 (2006).
13. Antoniali, G., Malfatti, M. C. & Tell, G. Unveiling the non-repair face of the Base Excision Repair pathway in RNA processing: A missing link between DNA repair and gene expression? *DNA Repair* **56**, 65–74 (2017).
14. Clark, D. W., Phang, T., Edwards, M. G., Geraci, M. W. & Gillespie, M. N. Promoter G-quadruplex sequences are targets for base oxidation and strand cleavage during hypoxia-induced transcription. *Free Radic. Biol. Med.* **53**, 51–59 (2012).
15. Park, J. W. *et al.* 8-OxoG in GC-rich Sp1 binding sites enhances gene transcription in adipose tissue of juvenile mice. *Sci. Rep.* **9**, 15618 (2019).
16. Fleming, A. M., Zhou, J., Wallace, S. S. & Burrows, C. J. A Role for the Fifth G-Track in G-Quadruplex Forming Oncogene Promoter Sequences during Oxidative Stress: Do These “Spare Tires” Have an Evolved Function? *ACS Cent. Sci.* **1**, 226–233 (2015).
17. Omega, C. A., Fleming, A. M. & Burrows, C. J. The Fifth Domain in the G-Quadruplex-Forming Sequence of the Human NEIL3 Promoter Locks DNA Folding in Response to Oxidative Damage. *Biochemistry* **57**, 2958–2970 (2018).
18. Zhu, J., Fleming, A. M. & Burrows, C. J. The RAD17 Promoter Sequence Contains a Potential Tail-Dependent G-Quadruplex That Downregulates Gene Expression upon Oxidative Modification. *ACS Chem. Biol.* **13**, 2577–2584 (2018).
19. Bielskutė, S., Plavec, J. & Podbevšek, P. Impact of Oxidative Lesions on the Human Telomeric G-Quadruplex. *J. Am. Chem. Soc.* **141**, 2594–2603 (2019).

20. Virgilio, A. *et al.* The oxidative damage to the human telomere: effects of 5-hydroxymethyl-2'-deoxyuridine on telomeric G-quadruplex structures. *Org. Biomol. Chem.* **13**, 7421–7429 (2015).
21. Bielskutè, S., Plavec, J. & Podbevšek, P. Oxidative lesions modulate G-quadruplex stability and structure in the human BCL2 promoter. *Nucleic Acids Res.* **49**, 2346–2356 (2021).
22. Cheong, V. V., Heddi, B., Lech, C. J. & Phan, A. T. Xanthine and 8-oxoguanine in G-quadruplexes: formation of a G·G·X·O tetrad. *Nucleic Acids Res.* **43**, 10506–10514 (2015).
23. Kypr, J., Kejnovská, I., Renčiuk, D. & Vorlíčková, M. Circular dichroism and conformational polymorphism of DNA. *Nucleic Acids Res.* **37**, 1713–1725 (2009).
24. Vorlíčková, M., Kejnovská, I., Bednářová, K., Renčiuk, D. & Kypr, J. Circular Dichroism Spectroscopy of DNA: From Duplexes to Quadruplexes. *Chirality* **24**, 691–698 (2012).
25. del Villar-Guerra, R., Trent, J. O. & Chaires, J. B. G-Quadruplex Secondary Structure Obtained from Circular Dichroism Spectroscopy. *Angew. Chem. Int. Ed.* **57**, 7171–7175 (2018).
26. Wang, K.-B. *et al.* Oxidative Damage Induces a Vacancy G-Quadruplex That Binds Guanine Metabolites: Solution Structure of a cGMP Fill-in Vacancy G-Quadruplex in the Oxidized BLM Gene Promoter. *J. Am. Chem. Soc.* **144**, 6361–6372 (2022).
27. Xiao, X. *et al.* APP, PSEN1, and PSEN2 Variants in Alzheimer's Disease: Systematic Re-evaluation According to ACMG Guidelines. *Front. Aging Neurosci.* **13**, (2021).
28. Renbaum, P. *et al.* Egr-1 upregulates the Alzheimer's disease presenilin-2 gene in neuronal cells. *Gene* **318**, 113–124 (2003).
29. Kikin, O., D'Antonio, L. & Bagga, P. S. QGRS Mapper: a web-based server for predicting G-quadruplexes in nucleotide sequences. *Nucleic Acids Res.* **34**, W676–W682 (2006).
30. Breen, A. P. & Murphy, J. A. Reactions of oxyl radicals with DNA. *Free Radic. Biol. Med.* **18**, 1033–1077 (1995).

31. Burrows, C. J. & Muller, J. G. Oxidative Nucleobase Modifications Leading to Strand Scission. *Chem. Rev.* **98**, 1109–1152 (1998).
32. Lloyd, D. R. & Phillips, D. H. Oxidative DNA damage mediated by copper(II), iron(II) and nickel(II) Fenton reactions: evidence for site-specific mechanisms in the formation of double-strand breaks, 8-hydroxydeoxyguanosine and putative intrastrand cross-links. *Mutat. Res. Mol. Mech. Mutagen.* **424**, 23–36 (1999).
33. Maniatis, T., Jeffrey, A. & Van deSande, H. Chain length determination of small double- and single-stranded DNA molecules by polyacrylamide gel electrophoresis. *Biochemistry* **14**, 3787–3794 (1975).
34. Gorini, F., Ambrosio, S., Lania, L., Majello, B. & Amente, S. The Intertwined Role of 8-oxodG and G4 in Transcription Regulation. *Int. J. Mol. Sci.* **24**, 2031 (2023).
35. Müller, N. & Khobta, A. Regulation of GC box activity by 8-oxoguanine. *Redox Biol.* **43**, 101997 (2021).
36. Hahm, J. Y., Park, J., Jang, E.-S. & Chi, S. W. 8-Oxoguanine: from oxidative damage to epigenetic and epitranscriptional modification. *Exp. Mol. Med.* **54**, 1626–1642 (2022).
37. Brown, R. V. *et al.* The Consequences of Overlapping G-Quadruplexes and i-Motifs in the Platelet-Derived Growth Factor Receptor β Core Promoter Nuclease Hypersensitive Element Can Explain the Unexpected Effects of Mutations and Provide Opportunities for Selective Targeting of Both Structures by Small Molecules To Downregulate Gene Expression. *J. Am. Chem. Soc.* **139**, 7456–7475 (2017).
38. Zhou, J., Fleming, A. M., Averill, A. M., Burrows, C. J. & Wallace, S. S. The NEIL glycosylases remove oxidized guanine lesions from telomeric and promoter quadruplex DNA structures. *Nucleic Acids Res.* **43**, 4039–4054 (2015).
39. Delabio, R. *et al.* PSEN1 and PSEN2 Gene Expression in Alzheimer's Disease Brain: A New Approach. *J. Alzheimers Dis.* **42**, 757–760 (2014).

40. Gray, D. M. *et al.* Measured and calculated CD spectra of G-quartets stacked with the same or opposite polarities. *Chirality* **20**, 431–440 (2008).
41. Henle, E. S., Luo, Y., Gassmann, W. & Linn, S. Oxidative Damage to DNA Constituents by Iron-mediated Fenton Reactions. *J. Biol. Chem.* **271**, 21177–21186 (1996).
42. Bruskov, V. I., Malakhova, L. V., Masalimov, Z. K. & Chernikov, A. V. Heat-induced formation of reactive oxygen species and 8-oxoguanine, a biomarker of damage to DNA. *Nucleic Acids Res.* **30**, 1354–1363 (2002).
43. Kim, J.-E., Choi, S., Yoo, J.-A. & Chung, M.-H. 8-Oxoguanine induces intramolecular DNA damage but free 8-oxoguanine protects intermolecular DNA from oxidative stress. *FEBS Lett.* **556**, 104–110 (2004).

Summary and Future Work

The work presented in this thesis is the first time a DNA crystal was soaked in hydrogen peroxide to attempt to induce Fenton-based oxidative damage and identify preferred sites where this could occur. To find suitable systems, the methodology used to obtain ordered metal sites had to be considered. Cocrystallisation and soaking are two methods mainly used to get transition metals into the crystals. The work presented in Chapter 2 demonstrates how the DNA sequence d(GCATGCT), known to crystallise as non-canonical quadruplexes stabilised by guanine-cytosine tetrads, adopts two distinct structures when crystallised using these two methods. The ordered transition-metal ions found in the four reported structures were nickel (II) ions and cobalt (II) ions, two transition metals which may act as models for iron (II) or copper (II) and produce reactive oxygen species that can DNA damage. However, availability of the transition metals to the solvent channels, and hence, the hydrogen peroxide, is essential in this project. Changes in the structure and crosslinking are common with cocrystallisation, and the transition metal would not be available to the solvent channels, thus, cocrystallisation was found to be unsuitable. This was not observed with soaking and this method was chosen to further investigate oxidative damage in DNA using crystals.

To further examine transition-metal binding preferences, additional transition metals were investigated, especially transition metals which have been associated with the formation of reactive oxygen species (ROS) in the presence of hydrogen peroxide and associated to oxidative damage in DNA. Copper and iron both meet the criteria, binding preferences for these two transition metals were therefore examined. Copper (II) binding preferences were investigated in Chapter 3. Crystals of the DNA sequence d(GCATGCT) were soaked in copper (II) chloride. Copper has been determined to bind primarily at the N7 position of guanines, a position which has been described to help “deliver hydroxyl radical via binding”¹. The hypothesis that oxidative damage occurs where the copper is bound was consequently considered. Two out of the four guanine bases exposed to the solvent channels in a single quadruplex were reported to be coordinated to copper (II) ions. Furthermore, providing oxidative damage occur where the copper is bound, the oxidative damage marker, 8-

oxoguanine, was incorporated to the d(GCATGCT) sequence at the position where copper (II) ions were found to coordinate previously. The crystal structure of d(oxoGCATGCT) was obtained, and the influence of 8-oxo-G on copper binding preferences was described. Copper (II) ions could no longer coordinate directly to the N7 position of oxidised guanines but were found to bind to the N7 positions of the unbound guanines from the native structure. Hence, a novel mechanism of oxidative damage progression was hypothesised.

Iron (II) ions binding preferences were investigated in Chapter 4 and the first DNA structure containing iron (II) was reported. The DNA sequence d(CGCGCG), known to crystallise as a Z-DNA duplex, was previously used to investigate transition-metal binding preferences. The soaking experiment was performed in an anoxic environment to prevent air oxidation of the iron (II) ions present in solution and in the crystal. Four iron (II) ions were reported to be coordinated in a single duplex, as opposed to the six copper (II) ions identified after copper soaking in the same system. The variation in binding preferences has been hypothesised as an explanation as to why oxidative damage differs between copper-catalysed Fenton and iron-catalysed Fenton.

DNA crystals were soaked in hydrogen peroxide for the first time in Chapter 5. Copper-soaked crystals of the Z-DNA sequence d(CGCGCG) were used for this experiment and were consecutively soaked in hydrogen peroxide to attempt to induce oxidative damage and identify potential hotspots in the structure. Hydrogen peroxide molecules were found coordinated to the copper (II) ions in all the structures when the data was collected at 100 K. Furthermore, copper centres disappeared as the reaction occurred, an indication that the copper (II) ions unbind from the DNA as a consequence of repeated peroxide exposure, potentially as a consequence of an oxidation state change.

Chapter 6 describes how a G-quadruplex forming sequence, PSEN2, a gene implicated in Alzheimer's disease, was found to undergo conformational changes in an oxidative stress environment. The structural effects of the presence of more than four tracts of guanines in the sequence were investigated and the incorporation of base substitutions to suppress quadruplex

formation and mimic oxidative damage gave insight into the potential equilibrium between all the potential G-quadruplex topologies in this system and how oxidative damage can influence topological changes. Conformational changes in the promoter region of this gene could act as a regulator for its expression and this mechanism could be correlated to the increase in oxidative damage found in patient suffering from Alzheimer's disease.

In this work, copper (II) and iron (II) binding preferences were investigated in two different DNA systems. However, nucleic acids are flexible and can form many 3-dimensional structures, and the results from these two systems could hardly be used to describe all the possible binding poses these two transition metals could have when binding to nucleic acids. Little previous work to identify the binding preferences for copper (II) and iron (II) have been undertaken. The work reported here could be used as a starting point to examine more complex nucleic acid systems including G-quadruplexes and other non-canonical nucleic acid structures. Moreover, transition metal ions may influence the structure of DNA when bound. Whilst cross-linking between DNA strands within a crystal is a common occurrence, metals are known to also influence DNA topology and conformation. A library of sequences could be produced to systematically explore how the binding of Cu (II) and Fe (II) can influence DNA structure as a function of sequence. This could then be extended to understand how transition metals might promote or prevent DNA structure formation could help further analyse how and where Fenton-based oxidative damage occurs on the DNA strand.

X-ray crystallography and circular dichroism spectroscopy are techniques that will provide structural information about transition-metal binding preferences. However, other techniques could be used to characterise transition-metal – DNA binding. Isothermal titration calorimetry (ITC) can be utilised to characterise DNA – transition-metal interactions and determine binding affinity. This technique can also provide information on stoichiometry and reveal how many transition metals are bound to a single DNA molecule. UV-vis spectroscopy can be applied as another approach to detect and monitor transition-metal coordination in solution.

Whilst it was possible to visualise peroxide bound to metal centres after soaking, evidence of oxidative damage in DNA was not present and so the method should be explored further. Time-resolved X-ray crystallography is a powerful technique to detect molecular changes following a reaction², but the diffusion of the reactant into large crystals can be relatively slow. The emergence of micro-focused X-ray beams at synchrotrons or XFEL sources means microcrystals can be used instead. How the hydrogen peroxide is delivered in the sample could also be reimagined, as methods such as the mix-and-inject serial crystallography method use sample delivery systems which include a mixing step before injecting the sample into the X-ray beam³. Furthermore, access to X-ray absorption fine structure (XAFS) spectroscopy or X-ray Fluorescence (XRF) could help identify transition-metal signatures to confirm the transition metal has soaked successfully into the crystal. New methods could also be integrated into the data analysis pipeline to confirm the presence of oxidative damage in DNA and transition-metal oxidation state. PanDDa is a software which could be used to detect and identify the weak signal associated with guanine oxidation, while the Spatially resolved anomalous dispersion (SpReAD) refinement method can help determine transition-metals oxidation state, as higher oxidation states result in absorption edges to shift to a higher energy⁴. Furthermore, crystals containing DNA treated with Fenton reagents could be dissolved and analysed using Mass spectrometry and HPLC to identify damage lesions and correlate this with appearance in the crystal structure, to validate weak electron density peaks observed at potential oxidation sites.

References

1. Fleming, A. M., Muller, J. G., Ji, I. & Burrows, C. J. Characterization of 2'-deoxyguanosine oxidation products observed in the Fenton-like system Cu(II)/H₂O₂/reductant in nucleoside and oligodeoxynucleotide contexts. *Org. Biomol. Chem.* **9**, 3338–3348 (2011).
2. Brändén, G. & Neutze, R. Advances and challenges in time-resolved macromolecular crystallography. *Science* **373**, eaba0954 (2021).
3. Olmos, J. L. *et al.* Enzyme intermediates captured “on the fly” by mix-and-inject serial crystallography. *BMC Biol.* **16**, 59 (2018).
4. Lennartz, F., Jeoung, J.-H., Ruenger, S., Dobbek, H. & Weiss, M. S. Determining the oxidation state of elements by X-ray crystallography. *Acta Crystallogr. Sect. Struct. Biol.* **78**, 238–247 (2022).

Supplementary Information

Table A1: Local base-pair parameters for 8OE3.

bp	Shear	Stretch	Stagger	Buckle	Propeller	Opening
1 G-C	-3.43	8.39	-0.48	-3.64	38.95	173.85
2 C-G	3.21	8.41	-0.24	9.40	35.90	176.42
3 G-C	-3.08	8.45	-0.61	-12.71	35.02	174.11
4 C-G	3.76	8.22	-0.03	7.65	35.74	176.90
~~~~~						
ave.	0.12	8.37	-0.34	0.18	36.40	175.32
s.d.	3.90	0.10	0.26	10.35	1.74	1.56

**Table A2: Local base-pair step parameters for 8OE3.**

step	Shift	Slide	Rise	Tilt	Roll	Twist
1 GC/GC	0.24	0.64	2.95	-1.41	5.33	30.52
2 CG/CG	----	----	----	----	----	----
3 GC/GC	0.14	0.57	3.02	-2.58	5.03	33.82

**Table A3: Local base-pair helical parameters for 8OE3.**

step	X-disp	Y-disp	h-Rise	Incl.	Tip	h-Twist
1 GC/GC	0.26	-0.70	3.01	10.02	2.65	31.00
2 CG/CG	----	----	----	----	----	----
3 GC/GC	0.23	-0.61	3.05	8.58	4.40	34.27

**Table A4: Local base-pair parameters for 8ASO.**

bp	Shear	Stretch	Stagger	Buckle	Propeller	Opening
1 G-C	-3.46	9.07	-0.36	-3.31	38.66	177.38
2 C-G	3.22	8.40	-0.22	8.91	36.13	176.23
3 A-A	0.06	-5.56	-1.78	-0.33	-37.35	176.71
4 G-C	-3.77	8.17	-0.08	-10.21	35.67	176.25
5 C-G	3.09	7.58	-0.87	13.74	24.56	175.15
~~~~~						
ave.	-0.17	5.53	-0.66	1.76	19.53	176.34
s.d.	3.39	6.22	0.69	9.59	32.26	0.81

Table A5: Local base-pair step parameters for 8ASO.

step	Shift	Slide	Rise	Tilt	Roll	Twist
1 GC/GC	0.23	0.47	2.88	-0.59	5.29	28.80
2 CA/AG	2.61	-0.17	5.74	1.42	2.41	-12.59
3 AG/CA	-2.57	-0.21	5.97	-1.55	-0.57	-12.27
4 GC/GC	-0.09	0.77	3.15	9.25	3.89	36.99
~~~~~						
ave.	0.05	0.21	4.44	2.13	2.75	10.24
s.d.	2.12	0.48	1.64	4.90	2.51	26.38

**Table A6: Local base-pair helical parameters for 8ASO.**

step	X-disp	Y-disp	h-Rise	Incl.	Tip	h-Twist
1 GC/GC	-0.09	-0.57	2.92	10.52	1.17	29.28
2 CA/AG	-3.89	14.38	5.35	-10.80	6.35	-12.89
3 AG/CA	2.23	-15.28	5.59	2.65	-7.20	-12.38
4 GC/GC	0.70	1.27	3.10	6.00	-14.26	38.28
~~~~~						
ave.	-0.27	-0.05	4.24	2.09	-3.48	10.57
s.d.	2.60	12.14	1.42	9.18	9.10	27.05

Table A7: Local base-pair parameters for 8ASM.

bp	Shear	Stretch	Stagger	Buckle	Propeller	Opening
1 G-C	-3.55	8.94	-0.25	-3.85	37.51	178.77
2 C-G	3.30	8.44	-0.34	8.36	36.60	175.76
3 G-C	-2.90	8.63	-0.68	-6.94	36.58	173.73
4 C-G	3.89	8.30	-0.08	9.47	34.29	177.75
~~~~~						
ave.	0.18	8.58	-0.34	1.76	36.25	176.50
s.d.	3.95	0.28	0.25	8.37	1.37	2.23

**Table A8: Local base-pair step parameters for 8ASM.**

step	Shift	Slide	Rise	Tilt	Roll	Twist
1 GC/GC	0.22	0.59	2.96	-0.13	4.49	30.10
2 CG/CG	----	----	----	----	----	----
3 GC/GC	0.18	0.47	2.97	-2.45	5.01	31.65

**Table A9: Local base-pair helical parameters for 8ASM.**

step	X-disp	Y-disp	h-Rise	Incl.	Tip	h-Twist
1 GC/GC	0.31	-0.44	3.02	8.59	0.24	30.43
2 CG/CG	----	----	----	----	----	----
3 GC/GC	0.03	-0.72	2.99	9.10	4.45	32.13

**Table A10: Local base-pair parameters for 8BAE.**

bp	Shear	Stretch	Stagger	Buckle	Propeller	Opening
1 G-C	-3.34	9.02	-0.39	-2.93	39.10	177.98
2 C-G	3.21	8.46	-0.27	8.69	36.20	176.12
3 G-C	-3.22	8.57	-0.57	-6.84	36.72	173.56
4 C-G	3.83	8.20	-0.09	9.44	35.48	176.59
~~~~~						
ave.	0.12	8.56	-0.33	2.09	36.88	176.06
s.d.	3.93	0.34	0.20	8.22	1.57	1.85

Table A11: Local base-pair step parameters for 8BAE.

step	Shift	Slide	Rise	Tilt	Roll	Twist
1 GC/GC	0.26	0.48	2.88	-0.79	6.15	28.83
2 CG/CG	----	----	----	----	----	----
3 GC/GC	0.18	0.83	2.91	-2.08	7.24	34.62

Table A12: Local base-pair helical parameters for 8BAE.

step	X-disp	Y-disp	h-Rise	Incl.	Tip	h-Twist
1 GC/GC	-0.23	-0.66	2.91	12.18	1.57	29.47
2 CG/CG	----	----	----	----	----	----
3 GC/GC	0.39	-0.57	3.00	11.98	3.44	35.41

Table A13: Local base-pair parameters for 8BAF.

bp	Shear	Stretch	Stagger	Buckle	Propeller	Opening
1 g-C	-3.30	8.58	-0.35	-6.62	34.57	175.87
2 C-G	3.36	8.39	-0.22	9.67	36.29	177.18
3 g-C	-2.99	8.42	-0.58	-13.71	32.87	176.69
4 C-G	3.85	8.27	-0.10	8.38	36.11	178.31
~~~~~						
ave.	0.23	8.42	-0.31	-0.57	34.96	177.01
s.d.	3.91	0.13	0.20	11.46	1.59	1.02

**Table A14: Local base-pair step parameters for 8BAF.**

step	Shift	Slide	Rise	Tilt	Roll	Twist
1 gC/GC	0.22	0.63	2.95	-2.11	6.36	31.91
2 Cg/CG	----	----	----	----	----	----
3 gC/GC	0.18	0.44	2.99	-3.07	5.97	34.61

**Table A15: Local base-pair helical parameters for 8BAF.**

step	X-disp	Y-disp	h-Rise	Incl.	Tip	h-Twist
1 gC/GC	0.11	-0.72	3.00	11.40	3.79	32.59
2 Cg/CG	----	----	----	----	----	----
3 gC/GC	-0.09	-0.71	2.99	9.92	5.10	35.24

**Table A16: Local base-pair parameters for 8BAG.**

bp	Shear	Stretch	Stagger	Buckle	Propeller	Opening
1 g-C	-3.26	8.27	-0.53	-9.15	36.35	173.58
2 C-G	3.27	8.41	-0.34	8.38	37.49	176.04
3 g-C	-3.13	8.06	-0.61	-18.32	28.17	177.42
4 C-G	3.82	8.22	-0.16	7.53	37.13	177.06
~~~~~						
ave.	0.18	8.24	-0.41	-2.89	34.78	176.03
s.d.	3.90	0.14	0.20	13.07	4.44	1.73

Table A17: Local base-pair step parameters for 8BAG.

step	Shift	Slide	Rise	Tilt	Roll	Twist
1 gC/GC	0.27	0.77	2.95	-1.91	6.69	32.40
2 Cg/CG	----	----	----	----	----	----
3 gC/GC	0.13	0.53	3.09	-4.80	4.76	36.21

Table A18: Local base-pair helical parameters for 8BAG.

step	X-disp	Y-disp	h-Rise	Incl.	Tip	h-Twist
1 gC/GC	0.31	-0.77	3.03	11.83	3.38	33.12
2 Cg/CG	----	----	----	----	----	----
3 gC/GC	0.23	-0.82	3.09	7.58	7.63	36.82

Table A19: Local base-pair parameters for 8OEX.

bp	Shear	Stretch	Stagger	Buckle	Propeller	Opening
1 C-G	-0.22	-0.13	0.08	1.37	-2.58	2.69
2 G-C	0.25	-0.16	0.12	-3.04	3.31	1.82
3 C-G	-0.32	-0.15	0.07	2.58	7.10	2.43
4 G-C	0.26	-0.18	0.07	-3.71	1.18	2.45
5 C-G	-0.23	-0.16	-0.05	5.65	-7.51	1.14
6 G-C	0.49	-0.04	0.33	-0.57	0.50	-1.20
~~~~~						
ave.	0.04	-0.14	0.10	0.38	0.33	1.56
s.d.	0.34	0.05	0.13	3.55	5.01	1.46

**Table A20: Local base-pair step parameters for 8OEX.**

step	Shift	Slide	Rise	Tilt	Roll	Twist
1 CG/CG	0.26	5.43	3.54	2.38	-2.07	-9.57
2 GC/GC	0.67	-0.30	3.20	0.97	-6.55	-54.75
3 CG/CG	-0.45	5.54	3.63	-2.03	2.94	-7.05
4 GC/GC	-0.73	-1.11	3.28	1.97	-1.68	-51.57
5 CG/CG	-0.05	5.41	3.69	0.22	1.13	-5.31
~~~~~						
ave.	-0.06	2.99	3.47	0.70	-1.25	-25.65
s.d.	0.55	3.39	0.22	1.74	3.61	25.18

Table A21: Local base-pair helical parameters for 8OEX.

step	X-disp	Y-disp	h-Rise	Incl.	Tip	h-Twist
1 CG/CG	-26.17	6.96	4.42	11.97	13.79	-10.08
2 GC/GC	0.70	0.78	3.13	7.09	1.06	-55.12
3 CG/CG	-46.87	-7.88	1.06	-22.17	-15.25	-7.90
4 GC/GC	1.38	-0.70	3.27	1.93	2.26	-51.63
5 CG/CG	-63.86	0.76	2.48	-11.99	2.32	-5.43
~~~~~						
ave.	-26.96	-0.02	2.87	-2.63	0.83	-26.03
s.d.	28.84	5.30	1.23	14.12	10.38	25.04

**Table A22: Local base-pair parameters for 8OEY.**

bp	Shear	Stretch	Stagger	Buckle	Propeller	Opening
1 C-G	-0.30	-0.17	0.17	1.88	6.74	3.51
2 G-C	0.35	-0.20	0.14	-0.44	6.09	2.84
3 C-G	-0.28	-0.11	0.04	8.91	-1.80	1.46
4 G-C	0.32	-0.12	-0.10	-9.76	-3.77	-1.34
5 C-G	-0.06	-0.22	-0.19	3.25	-0.50	3.18
6 G-C	0.13	-0.19	0.16	-4.40	-0.74	1.79
~~~~~						
ave.	0.03	-0.17	0.04	-0.09	1.00	1.91
s.d.	0.29	0.05	0.15	6.45	4.35	1.78

Table A23: Local base-pair step parameters for 8OEY.

step	Shift	Slide	Rise	Tilt	Roll	Twist
1 CG/CG	0.06	5.60	3.56	1.83	5.41	-12.26
2 GC/GC	-0.18	-0.37	3.20	0.25	-1.96	-56.26
3 CG/CG	-0.12	5.39	3.80	0.73	-6.67	-4.14
4 GC/GC	0.06	-1.28	3.15	2.52	-6.73	-46.67
5 CG/CG	-0.07	5.57	3.63	-3.57	-0.10	-10.64
~~~~~						
ave.	-0.05	2.98	3.47	0.35	-2.01	-25.99
s.d.	0.11	3.49	0.28	2.36	5.07	23.70

**Table A24: Local base-pair helical parameters for 8OEY.**

step	X-disp	Y-disp	h-Rise	Incl.	Tip	h-Twist
1 CG/CG	-27.83	1.61	0.99	-23.73	8.01	-13.52
2 GC/GC	0.50	-0.18	3.19	2.08	0.27	-56.30
3 CG/CG	2.58	3.66	6.57	57.99	6.32	-7.88
4 GC/GC	2.12	0.27	2.94	8.44	3.15	-47.19
5 CG/CG	-28.32	-6.25	3.47	0.51	-18.56	-11.22
~~~~~						
ave.	-10.19	-0.18	3.43	9.06	-0.16	-27.22
s.d.	16.35	3.71	2.01	29.96	10.71	22.71

Table A25: Local base-pair parameters for 8OEZ.

bp	Shear	Stretch	Stagger	Buckle	Propeller	Opening
1 C-G	-0.24	-0.27	0.28	3.82	4.05	2.50
2 G-C	0.22	-0.09	0.04	-5.92	0.62	1.92
3 C-G	-0.38	-0.12	-0.08	5.61	3.55	1.51
4 G-C	0.25	-0.03	0.15	-3.60	0.37	0.59
5 C-G	-0.26	-0.05	0.10	2.48	-2.33	0.82
6 G-C	0.34	-0.03	0.27	-5.93	2.86	0.38
~~~~~						
ave.	-0.01	-0.10	0.13	-0.59	1.52	1.29
s.d.	0.31	0.09	0.14	5.16	2.42	0.83

**Table A26: Local base-pair step parameters for 8OEZ.**

step	Shift	Slide	Rise	Tilt	Roll	Twist
1 CG/CG	0.05	5.64	3.71	1.28	0.53	-12.13
2 GC/GC	0.43	-0.52	3.13	0.66	-5.68	-51.16
3 CG/CG	-0.15	5.37	3.66	-2.29	-0.07	-6.64
4 GC/GC	-0.13	-1.19	3.31	1.83	-3.58	-47.66
5 CG/CG	-0.06	5.57	3.79	-1.44	1.04	-11.87
~~~~~						
ave.	0.03	2.97	3.52	0.01	-1.55	-25.89
s.d.	0.24	3.50	0.28	1.79	2.93	21.62

Table A27: Local base-pair helical parameters for 8OEZ.

step	X-disp	Y-disp	h-Rise	Incl.	Tip	h-Twist
1 CG/CG	-27.22	2.00	3.44	-2.49	6.02	-12.21
2 GC/GC	0.97	0.54	3.06	6.56	0.76	-51.46
3 CG/CG	-43.50	-11.00	3.47	0.60	-19.04	-7.02
4 GC/GC	1.76	-0.02	3.22	4.41	2.26	-47.82
5 CG/CG	-28.11	-2.30	3.26	-5.01	-6.93	-12.00
~~~~~						
ave.	-19.22	-2.16	3.29	0.82	-3.38	-26.10
s.d.	19.88	5.18	0.17	4.76	9.94	21.63

**Table A28: Local base-pair parameters for strand A/B in 8OE7.**

bp	Shear	Stretch	Stagger	Buckle	Propeller	Opening
1 C-G	-0.20	-0.10	0.08	2.80	-0.31	2.20
2 G-C	0.25	-0.14	0.05	-2.38	2.73	1.07
3 C-G	-0.27	-0.16	-0.06	2.89	6.99	2.94
4 G-C	0.26	-0.22	0.01	-2.41	2.90	2.79
5 C-G	-0.06	-0.29	0.20	1.48	-3.66	2.64
6 G-C	0.24	-0.14	0.33	-2.71	2.11	1.83
~~~~~						
ave.	0.04	-0.17	0.10	-0.05	1.79	2.24
s.d.	0.24	0.07	0.14	2.73	3.56	0.70

Table A29: Local base-pair step parameters for strand A/B in 8OE7.

step	Shift	Slide	Rise	Tilt	Roll	Twist
1 CG/CG	0.05	5.64	3.71	1.28	0.53	-12.13
2 GC/GC	0.43	-0.52	3.13	0.66	-5.68	-51.16
3 CG/CG	-0.15	5.37	3.66	-2.29	-0.07	-6.64
4 GC/GC	-0.13	-1.19	3.31	1.83	-3.58	-47.66
5 CG/CG	-0.06	5.57	3.79	-1.44	1.04	-11.87
~~~~~						
ave.	0.03	2.97	3.52	0.01	-1.55	-25.89
s.d.	0.24	3.50	0.28	1.79	2.93	21.62

**Table A30: Local base-pair helical parameters for strand A/B in 8OE7.**

step	X-disp	Y-disp	h-Rise	Incl.	Tip	h-Twist
1 CG/CG	-27.68	6.07	4.29	10.38	11.63	-9.92
2 GC/GC	0.79	0.77	3.20	7.03	0.63	-54.38
3 CG/CG	-38.75	-10.46	3.09	-2.25	-19.68	-8.38
4 GC/GC	1.60	-0.58	3.37	2.08	-0.55	-50.24
5 CG/CG	-44.90	2.18	1.51	-20.63	4.44	-8.27
~~~~~						
ave.	-21.79	-0.41	3.09	-0.68	-0.70	-26.24
s.d.	21.87	6.15	1.01	12.14	11.62	23.86

Table A31: Local base-pair parameters for strand C/D in 8OE7.

bp	Shear	Stretch	Stagger	Buckle	Propeller	Opening
1 C-G	-0.23	-0.13	0.21	3.23	1.07	2.82
2 G-C	0.47	0.02	-0.22	-10.22	7.25	-0.05
3 C-G	-0.34	-0.15	-0.09	5.34	9.08	3.28
4 G-C	0.33	-0.13	-0.02	-5.35	1.70	1.92
5 C-G	-0.20	-0.20	0.08	4.10	-6.22	1.86
6 G-C	0.29	-0.17	0.14	-5.25	3.18	1.79
~~~~~						
ave.	0.05	-0.13	0.02	-1.36	2.68	1.94
s.d.	0.35	0.08	0.16	6.41	5.38	1.14

**Table A32: Local base-pair step parameters for strand C/D in 8OE7.**

step	Shift	Slide	Rise	Tilt	Roll	Twist
1 CG/CG	0.23	5.48	3.70	5.11	-3.56	-8.68
2 GC/GC	0.67	-0.13	3.00	-1.22	-4.60	-55.18
3 CG/CG	-0.43	5.53	3.63	-3.12	-2.46	-6.53
4 GC/GC	-0.48	-1.08	3.38	-0.26	0.22	-51.45
5 CG/CG	0.12	5.53	3.76	0.64	0.33	-6.98
~~~~~						
ave.	0.02	3.07	3.50	0.23	-2.01	-25.77
s.d.	0.49	3.37	0.31	3.06	2.22	25.20

Table A33: Local base-pair helical parameters for strand C/D in 8OE7.

step	X-disp	Y-disp	h-Rise	Incl.	Tip	h-Twist
1 CG/CG	-21.18	13.12	4.72	20.39	29.27	-10.68
2 GC/GC	0.39	0.66	3.00	4.96	-1.31	-55.37
3 CG/CG	-30.67	-17.00	4.70	19.38	-24.59	-7.64
4 GC/GC	1.23	-0.58	3.39	-0.25	-0.29	-51.45
5 CG/CG	-46.59	3.72	3.47	-2.71	5.24	-7.02
~~~~~						
ave.	-19.36	-0.02	3.86	8.35	1.66	-26.43
s.d.	20.53	10.90	0.80	10.89	19.22	24.71

**Table A34: Local base-pair parameters for strand E/F in 8OE7.**

bp	Shear	Stretch	Stagger	Buckle	Propeller	Opening
1 C-G	-0.26	-0.09	0.19	0.89	0.67	2.34
2 G-C	0.29	-0.16	-0.15	-6.57	5.03	1.25
3 C-G	-0.34	-0.16	-0.15	5.31	10.14	2.31
4 G-C	0.30	-0.21	-0.06	-3.55	0.96	2.71
5 C-G	-0.15	-0.25	-0.02	4.49	-3.51	3.72
6 G-C	0.26	-0.17	0.18	-2.54	2.05	2.44
~~~~~						
ave.	0.02	-0.17	-0.00	-0.33	2.56	2.46
s.d.	0.30	0.05	0.15	4.71	4.62	0.80

Table A35: Local base-pair step parameters for strand E/F in 8OE7.

step	Shift	Slide	Rise	Tilt	Roll	Twist
1 CG/CG	0.23	5.48	3.70	5.11	-3.56	-8.68
2 GC/GC	0.67	-0.13	3.00	-1.22	-4.60	-55.18
3 CG/CG	-0.43	5.53	3.63	-3.12	-2.46	-6.53
4 GC/GC	-0.48	-1.08	3.38	-0.26	0.22	-51.45
5 CG/CG	0.12	5.53	3.76	0.64	0.33	-6.98
~~~~~						
ave.	0.02	3.07	3.50	0.23	-2.01	-25.77
s.d.	0.49	3.37	0.31	3.06	2.22	25.20

**Table A36: Local base-pair helical parameters for strand E/F in 8OE7.**

step	X-disp	Y-disp	h-Rise	Incl.	Tip	h-Twist
1 CG/CG	-21.18	13.12	4.72	20.39	29.27	-10.68
2 GC/GC	0.39	0.66	3.00	4.96	-1.31	-55.37
3 CG/CG	-30.67	-17.00	4.70	19.38	-24.59	-7.64
4 GC/GC	1.23	-0.58	3.39	-0.25	-0.29	-51.45
5 CG/CG	-46.59	3.72	3.47	-2.71	5.24	-7.02
~~~~~						
ave.	-19.36	-0.02	3.86	8.35	1.66	-26.43
s.d.	20.53	10.90	0.80	10.89	19.22	24.71

Table A37: Local base-pair parameters for strand G/H in 8OE7.

bp	Shear	Stretch	Stagger	Buckle	Propeller	Opening
1 C-G	-0.24	-0.07	0.15	1.93	0.95	2.70
2 G-C	0.29	-0.14	0.15	-0.84	1.96	1.16
3 C-G	-0.27	-0.17	0.11	2.62	7.33	3.16
4 G-C	0.20	-0.21	0.15	-4.54	2.57	3.63
5 C-G	-0.25	-0.16	-0.10	3.67	-4.23	1.20
6 G-C	0.28	-0.14	0.26	-3.45	3.40	2.10
~~~~~						
ave.	0.00	-0.15	0.12	-0.10	1.99	2.32
s.d.	0.28	0.05	0.12	3.38	3.76	1.02

**Table A38: Local base-pair step parameters for strand G/H in 8OE7.**

step	Shift	Slide	Rise	Tilt	Roll	Twist
1 CG/CG	0.24	5.44	3.57	0.79	-1.57	-9.45
2 GC/GC	0.68	-0.45	3.28	0.70	-5.60	-54.34
3 CG/CG	-0.32	5.49	3.73	-2.29	0.35	-7.76
4 GC/GC	-0.50	-1.17	3.30	1.32	-1.87	-50.93
5 CG/CG	0.18	5.55	3.72	-0.49	2.62	-6.71
~~~~~						
ave.	0.06	2.97	3.52	0.01	-1.21	-25.84
s.d.	0.47	3.46	0.22	1.44	3.04	24.51

Table A39: Local base-pair helical parameters for strand G/H in 8OE7.

step	X-disp	Y-disp	h-Rise	Incl.	Tip	h-Twist
1 CG/CG	-28.56	3.42	4.38	9.43	4.72	-9.61
2 GC/GC	0.83	0.78	3.22	6.11	0.77	-54.61
3 CG/CG	-40.02	-9.41	3.24	-2.53	-16.47	-8.09
4 GC/GC	1.49	-0.49	3.27	2.17	1.53	-50.98
5 CG/CG	-51.78	-0.05	1.46	-21.31	-4.00	-7.22
~~~~~						
ave.	-23.61	-1.15	3.11	-1.23	-2.69	-26.10
s.d.	24.05	4.86	1.05	12.09	8.31	24.41

**Table A40: Local base-pair parameters for strand A/B in 8OE8.**

bp	Shear	Stretch	Stagger	Buckle	Propeller	Opening
1 C-G	-0.27	-0.14	0.23	2.12	1.56	2.44
2 G-C	0.21	-0.33	0.10	-5.70	2.71	5.49
3 C-G	-0.44	-0.13	0.01	4.75	4.97	2.72
4 G-C	0.77	-0.02	-0.01	-6.02	7.17	-1.75
5 C-G	-0.06	-0.33	-0.11	5.82	-1.91	4.83
6 G-C	0.34	-0.18	0.15	-3.43	3.32	1.99
~~~~~						
ave.	0.09	-0.19	0.06	-0.41	2.97	2.62
s.d.	0.44	0.12	0.12	5.30	3.09	2.56

Table A41: Local base-pair step parameters for strand A/B in 8OE8.

step	Shift	Slide	Rise	Tilt	Roll	Twist
1 CG/CG	0.35	5.57	3.74	1.49	0.60	-9.48
2 GC/GC	0.37	-0.58	3.16	0.87	-5.78	-53.92
3 CG/CG	-0.38	5.43	3.66	-0.48	-2.66	-4.23
4 GC/GC	-0.12	-0.90	3.28	-0.79	-0.12	-53.01
5 CG/CG	-0.11	5.57	3.79	0.20	1.87	-7.81
~~~~~						
ave.	0.02	3.02	3.52	0.26	-1.22	-25.69
s.d.	0.33	3.43	0.29	0.94	3.04	25.43

**Table A42: Local base-pair helical parameters for strand A/B in 8OE8.**

step	X-disp	Y-disp	h-Rise	Incl.	Tip	h-Twist
1 CG/CG	-34.58	5.34	3.28	-3.61	8.93	-9.61
2 GC/GC	0.98	0.46	3.08	6.35	0.96	-54.21
3 CG/CG	-30.57	-10.03	5.92	32.00	-5.81	-5.02
4 GC/GC	1.02	-0.18	3.27	0.13	-0.89	-53.02
5 CG/CG	-44.93	-0.22	2.39	-13.46	1.43	-8.03
~~~~~						
ave.	-21.62	-0.93	3.59	4.28	0.92	-25.98
s.d.	21.30	5.59	1.35	17.08	5.31	25.29

Table A43: Local base-pair parameters for strand C/D in 8OE8.

bp	Shear	Stretch	Stagger	Buckle	Propeller	Opening
1 C-G	-0.27	-0.15	0.18	1.86	1.51	2.16
2 G-C	0.28	-0.31	0.12	-4.69	1.61	5.06
3 C-G	-0.40	-0.12	0.02	4.98	5.77	2.65
4 G-C	0.79	-0.04	0.33	-1.15	2.71	-1.83
5 C-G	-0.10	-0.35	-0.12	6.23	-1.97	4.57
6 G-C	0.33	-0.19	0.12	-3.61	2.25	2.34
~~~~~						
ave.	0.10	-0.19	0.11	0.60	1.98	2.49
s.d.	0.44	0.12	0.15	4.50	2.49	2.44

**Table A44: Local base-pair step parameters for strand C/D in 8OE8.**

step	Shift	Slide	Rise	Tilt	Roll	Twist
1 CG/CG	0.31	5.59	3.66	1.32	0.58	-9.47
2 GC/GC	0.37	-0.59	3.19	1.36	-6.14	-53.78
3 CG/CG	-0.44	5.36	3.61	-2.87	-0.38	-4.59
4 GC/GC	-0.15	-1.02	3.33	2.31	-1.37	-53.42
5 CG/CG	-0.11	5.58	3.87	0.05	2.28	-7.57
~~~~~						
ave.	-0.01	2.98	3.53	0.44	-1.01	-25.77
s.d.	0.34	3.46	0.27	2.02	3.17	25.47

Table A45: Local base-pair helical parameters for strand C/D in 8OE8.

step	X-disp	Y-disp	h-Rise	Incl.	Tip	h-Twist
1 CG/CG	-34.75	4.74	3.24	-3.52	7.96	-9.58
2 GC/GC	1.01	0.49	3.11	6.77	1.50	-54.12
3 CG/CG	-53.87	-25.23	3.20	4.26	-31.98	-5.43
4 GC/GC	1.22	-0.02	3.31	1.52	2.57	-53.48
5 CG/CG	-46.80	-0.69	2.10	-16.79	0.37	-7.91
~~~~~						
ave.	-26.64	-4.14	2.99	-1.55	-3.91	-26.10
s.d.	26.24	11.98	0.50	9.34	15.96	25.33

**Table A46: Local base-pair parameters for strand E/F in 8OE8.**

bp	Shear	Stretch	Stagger	Buckle	Propeller	Opening
1 C-G	-0.22	-0.15	0.19	2.81	1.97	2.48
2 G-C	0.22	-0.30	0.32	-1.97	2.51	4.85
3 C-G	-0.36	-0.14	-0.03	5.43	5.84	2.93
4 G-C	0.57	-0.13	0.73	4.34	-0.88	-0.07
5 C-G	-0.12	-0.31	-0.15	5.75	-1.31	3.86
6 G-C	0.35	-0.19	0.12	-4.43	2.09	1.98
~~~~~						
ave.	0.07	-0.20	0.20	1.99	1.70	2.67
s.d.	0.36	0.08	0.31	4.22	2.60	1.69

Table A47: Local base-pair step parameters for strand E/F in 8OE8.

step	Shift	Slide	Rise	Tilt	Roll	Twist
1 CG/CG	0.29	5.58	3.62	0.34	0.51	-10.15
2 GC/GC	0.43	-0.60	3.28	2.58	-5.84	-53.24
3 CG/CG	-0.45	5.30	3.54	-6.11	1.19	-5.88
4 GC/GC	-0.23	-1.14	3.53	6.01	-2.02	-52.02
5 CG/CG	-0.02	5.60	3.83	-0.65	2.11	-7.63
~~~~~						
ave.	0.00	2.95	3.56	0.43	-0.81	-25.79
s.d.	0.36	3.49	0.20	4.47	3.21	24.56

**Table A48: Local base-pair helical parameters for strand E/F in 8OE8.**

step	X-disp	Y-disp	h-Rise	Incl.	Tip	h-Twist
1 CG/CG	-32.45	2.26	3.33	-2.89	1.90	-10.17
2 GC/GC	1.02	0.63	3.18	6.49	2.87	-53.60
3 CG/CG	-38.19	-16.88	1.37	-8.95	-45.81	-8.57
4 GC/GC	1.43	0.16	3.49	2.29	6.83	-52.38
5 CG/CG	-46.27	-1.94	2.19	-15.45	-4.76	-7.95
~~~~~						
ave.	-22.89	-3.15	2.71	-3.70	-7.80	-26.53
s.d.	22.56	7.82	0.90	8.75	21.66	24.17

Table A49: Local base-pair parameters for strand G/H in 8OE8.

bp	Shear	Stretch	Stagger	Buckle	Propeller	Opening
1 C-G	-0.26	-0.15	0.22	3.18	1.55	2.65
2 G-C	0.20	-0.30	0.23	-3.93	1.27	4.57
3 C-G	-0.36	-0.15	-0.04	5.22	5.57	2.70
4 G-C	0.47	-0.14	0.27	-1.32	3.34	0.52
5 C-G	-0.05	-0.32	-0.07	6.35	-1.91	4.69
6 G-C	0.34	-0.18	0.17	-2.97	2.65	1.62
~~~~~						
ave.	0.06	-0.21	0.13	1.09	2.08	2.79
s.d.	0.33	0.08	0.15	4.40	2.49	1.63

**Table A50: Local base-pair step parameters for strand G/H in 8OE8.**

step	Shift	Slide	Rise	Tilt	Roll	Twist
1 CG/CG	0.25	5.59	3.73	1.08	1.36	-10.09
2 GC/GC	0.41	-0.58	3.19	2.33	-6.38	-53.01
3 CG/CG	-0.37	5.42	3.67	-3.47	-0.61	-6.47
4 GC/GC	-0.23	-1.12	3.34	2.32	-1.37	-51.11
5 CG/CG	-0.05	5.57	3.81	-0.14	2.05	-8.03
~~~~~						
ave.	0.00	2.97	3.55	0.42	-0.99	-25.74
s.d.	0.33	3.50	0.27	2.41	3.32	24.07

Table A51: Local base-pair helical parameters for strand G/H in 8OE8.

step	X-disp	Y-disp	h-Rise	Incl.	Tip	h-Twist
1 CG/CG	-33.81	3.39	2.91	-7.64	6.07	-10.24
2 GC/GC	1.04	0.60	3.09	7.11	2.60	-53.41
3 CG/CG	-39.68	-16.89	3.50	4.97	-28.19	-7.37
4 GC/GC	1.40	-0.11	3.31	1.59	2.69	-51.18
5 CG/CG	-43.83	-0.73	2.31	-14.32	-0.96	-8.29
~~~~~						
ave.	-22.98	-2.75	3.02	-1.66	-3.56	-26.10
s.d.	22.37	8.06	0.46	9.05	13.99	23.95

**Table A52: Local base-pair parameters for 8OE9.**

bp	Shear	Stretch	Stagger	Buckle	Propeller	Opening
1 C-G	-0.20	-0.13	0.12	1.95	-1.34	3.16
2 G-C	0.26	-0.18	0.11	-3.58	2.70	2.20
3 C-G	-0.34	-0.13	0.05	3.70	6.30	3.10
4 G-C	0.25	-0.16	0.57	2.73	1.45	2.93
5 C-G	-0.22	-0.21	0.01	5.86	-5.48	2.01
6 G-C	0.17	-0.25	-0.06	-5.80	3.76	3.25
~~~~~						
ave.	-0.01	-0.18	0.13	0.81	1.23	2.78
s.d.	0.27	0.05	0.22	4.51	4.15	0.53

Table A53: Local base-pair step parameters for 8OE9.

step	Shift	Slide	Rise	Tilt	Roll	Twist
1 CG/CG	0.27	5.46	3.58	2.11	-1.69	-9.09
2 GC/GC	0.60	-0.40	3.19	0.91	-6.32	-54.65
3 CG/CG	-0.43	5.38	3.59	-5.14	2.33	-7.24
4 GC/GC	-0.58	-1.16	3.48	4.52	-0.74	-51.47
5 CG/CG	0.10	5.62	3.79	2.54	1.04	-7.65
~~~~~						
ave.	-0.01	2.98	3.53	0.99	-1.08	-26.02
s.d.	0.49	3.45	0.22	3.66	3.32	24.72

**Table A54: Local base-pair helical parameters for 8OE9.**

step	X-disp	Y-disp	h-Rise	Incl.	Tip	h-Twist
1 CG/CG	-28.66	7.05	4.35	10.38	12.94	-9.49
2 GC/GC	0.80	0.70	3.12	6.86	0.99	-54.99
3 CG/CG	-37.92	-12.10	1.22	-15.69	-34.59	-9.17
4 GC/GC	1.38	-0.33	3.50	0.85	5.19	-51.66
5 CG/CG	-42.76	8.20	2.82	-7.47	18.31	-8.13
~~~~~						
ave.	-21.43	0.70	3.00	-1.01	0.57	-26.69
s.d.	21.18	8.08	1.15	10.63	20.77	24.35

Table A55: Local base-pair parameters for strand A/B in 8OEA.

bp	Shear	Stretch	Stagger	Buckle	Propeller	Opening
1 C-G	-0.29	-0.34	0.20	-0.41	4.94	7.61
2 G-C	0.39	-0.26	-0.27	-8.27	8.96	1.15
3 C-G	0.19	-0.29	0.06	8.62	3.05	2.72
4 G-C	0.65	-0.13	0.24	-8.29	-4.43	-1.40
5 C-G	-0.50	-0.20	0.19	1.22	3.96	-0.82
6 G-C	0.24	-0.28	0.17	-3.66	2.93	7.23
~~~~~						
ave.	0.11	-0.25	0.10	-1.80	3.24	2.75
s.d.	0.43	0.07	0.19	6.43	4.36	3.90

**Table A56: Local base-pair step parameters for strand A/B in 8OEA.**

step	Shift	Slide	Rise	Tilt	Roll	Twist
1 CG/CG	0.08	5.73	3.59	3.89	1.05	-10.34
2 GC/GC	0.37	-0.34	2.93	-3.60	-6.77	-51.02
3 CG/CG	0.08	5.32	3.99	-5.57	-0.27	-5.70
4 GC/GC	-0.13	-1.09	3.27	3.66	-2.07	-53.47
5 CG/CG	0.44	5.61	3.82	-1.71	4.58	-8.89
~~~~~						
ave.	0.17	3.05	3.52	-0.66	-0.70	-25.89
s.d.	0.23	3.45	0.43	4.28	4.18	24.14

Table A57: Local base-pair helical parameters for strand A/B in 8OEA.

step	X-disp	Y-disp	h-Rise	Incl.	Tip	h-Twist
1 CG/CG	-31.23	6.36	2.78	-5.53	20.58	-11.10
2 GC/GC	0.80	0.20	2.88	7.81	-4.15	-51.56
3 CG/CG	-37.30	-20.15	3.08	2.14	-44.32	-7.97
4 GC/GC	1.34	0.08	3.23	2.30	4.06	-53.62
5 CG/CG	-38.16	0.10	0.89	-27.03	-10.09	-10.14
~~~~~						
ave.	-20.91	-2.68	2.57	-4.06	-6.78	-26.88
s.d.	20.24	10.13	0.96	13.69	23.95	23.51

**Table A58: Local base-pair parameters for strand C/D in 8OEA.**

bp	Shear	Stretch	Stagger	Buckle	Propeller	Opening
1 C-G	-0.77	0.08	0.26	4.82	3.31	-5.70
2 G-C	0.77	-0.08	-0.02	-7.36	1.37	-6.08
3 C-G	-0.80	0.01	-0.06	5.94	10.15	-3.60
4 G-C	0.07	0.01	0.08	-3.83	5.93	-0.06
5 C-G	0.11	-0.27	-0.08	8.12	-2.51	3.13
6 G-C	0.42	-0.03	0.47	1.66	1.91	-3.12
~~~~~						
ave.	-0.03	-0.05	0.11	1.56	3.36	-2.57
s.d.	0.64	0.12	0.22	6.02	4.32	3.53

Table A59: Local base-pair step parameters for strand C/D in 8OEA.

step	Shift	Slide	Rise	Tilt	Roll	Twist
1 CG/CG	0.07	5.42	3.79	2.12	-1.34	-5.83
2 GC/GC	0.75	-0.44	3.13	1.36	-5.94	-55.48
3 CG/CG	-0.41	5.58	3.74	-2.61	0.15	-9.87
4 GC/GC	-0.24	-1.00	3.26	0.30	-0.66	-46.54
5 CG/CG	-0.00	5.37	3.79	-0.75	2.64	-6.58
~~~~~						
ave.	0.04	2.99	3.54	0.08	-1.03	-24.86
s.d.	0.44	3.39	0.32	1.85	3.13	24.13

**Table A60: Local base-pair helical parameters for strand C/D in 8OEA.**

step	X-disp	Y-disp	h-Rise	Incl.	Tip	h-Twist
1 CG/CG	-40.59	13.84	4.60	12.47	19.69	-6.34
2 GC/GC	0.79	0.88	3.06	6.36	1.45	-55.79
3 CG/CG	-31.65	-7.53	3.43	-0.84	-14.82	-10.21
4 GC/GC	1.32	-0.27	3.25	0.84	0.38	-46.54
5 CG/CG	-51.35	-2.36	1.52	-21.76	-6.23	-7.13
~~~~~						
ave.	-24.29	0.91	3.17	-0.59	0.09	-25.20
s.d.	24.17	7.91	1.10	12.93	12.73	23.97

Table A61: Local base-pair parameters for 8OEB.

bp	Shear	Stretch	Stagger	Buckle	Propeller	Opening
1 C-G	-0.20	-0.13	0.12	1.95	-1.34	3.16
2 G-C	0.26	-0.18	0.11	-3.58	2.70	2.20
3 C-G	-0.34	-0.13	0.05	3.70	6.30	3.10
4 G-C	0.25	-0.16	0.57	2.73	1.45	2.93
5 C-G	-0.22	-0.21	0.01	5.86	-5.48	2.01
6 G-C	0.17	-0.25	-0.06	-5.80	3.76	3.25
~~~~~						
ave.	-0.01	-0.18	0.13	0.81	1.23	2.78
s.d.	0.27	0.05	0.22	4.51	4.15	0.53

**Table A62: Local base-pair step parameters for 8OEB.**

step	Shift	Slide	Rise	Tilt	Roll	Twist
1 CG/CG	0.27	5.58	3.60	1.88	-2.77	-10.10
2 GC/GC	0.50	-0.41	3.28	-0.54	-6.28	-53.98
3 CG/CG	-0.27	5.37	3.61	-3.14	0.42	-6.97
4 GC/GC	-0.40	-1.02	3.43	0.37	-0.92	-51.23
5 CG/CG	0.03	5.56	3.76	3.28	3.80	-8.90
~~~~~						
ave.	0.03	3.02	3.54	0.37	-1.15	-26.24
s.d.	0.37	3.41	0.19	2.45	3.74	24.12

Table A63: Local base-pair helical parameters for 8OEB.

step	X-disp	Y-disp	h-Rise	Incl.	Tip	h-Twist
1 CG/CG	-24.05	5.57	4.82	15.20	10.33	-10.64
2 GC/GC	0.83	0.51	3.22	6.89	-0.60	-54.32
3 CG/CG	-41.60	-12.49	2.89	-3.23	-24.29	-7.66
4 GC/GC	1.24	-0.43	3.42	1.06	0.42	-51.24
5 CG/CG	-36.75	4.97	1.21	-22.29	19.25	-10.22
~~~~~						
ave.	-20.06	-0.38	3.11	-0.47	1.03	-26.81
s.d.	20.30	7.27	1.30	14.02	16.30	23.75

**Table A64: Local base-pair parameters for 8OEC.**

bp	Shear	Stretch	Stagger	Buckle	Propeller	Opening
1 C-G	-0.28	-0.04	0.13	3.76	3.29	1.40
2 G-C	0.39	-0.23	0.18	-3.00	3.82	1.06
3 C-G	-0.46	-0.14	0.25	2.85	9.71	2.26
4 G-C	0.53	-0.21	0.20	-3.90	4.46	2.72
5 C-G	-0.22	-0.14	-0.08	8.12	-2.75	2.02
6 G-C	0.36	-0.21	0.21	-1.93	2.51	2.95
~~~~~						
ave.	0.05	-0.16	0.15	0.99	3.51	2.07
s.d.	0.42	0.07	0.12	4.70	3.99	0.74

Table A65: Local base-pair step parameters for 8OEC.

step	Shift	Slide	Rise	Tilt	Roll	Twist
1 CG/CG	0.17	5.53	3.69	1.24	1.41	-8.34
2 GC/GC	0.57	-0.51	3.28	0.45	-7.07	-55.21
3 CG/CG	-0.27	5.40	3.77	-1.69	0.67	-5.94
4 GC/GC	-0.43	-1.04	3.25	2.12	0.38	-52.21
5 CG/CG	0.03	5.54	3.91	-0.05	2.37	-6.64
~~~~~						
ave.	0.01	2.98	3.58	0.41	-0.45	-25.67
s.d.	0.39	3.43	0.30	1.43	3.78	25.64

**Table A66: Local base-pair helical parameters for 8OEC.**

step	X-disp	Y-disp	h-Rise	Incl.	Tip	h-Twist
1 CG/CG	-40.62	4.26	2.67	-9.51	8.40	-8.55
2 GC/GC	0.95	0.63	3.20	7.60	0.48	-55.63
3 CG/CG	-53.23	-11.12	2.96	-6.24	-15.85	-6.21
4 GC/GC	1.16	-0.35	3.27	-0.43	2.40	-52.25
5 CG/CG	-53.08	0.05	1.83	-19.62	-0.42	-7.05
~~~~~						
ave.	-28.96	-1.30	2.78	-5.64	-1.00	-25.94
s.d.	27.88	5.78	0.59	10.16	8.99	25.60

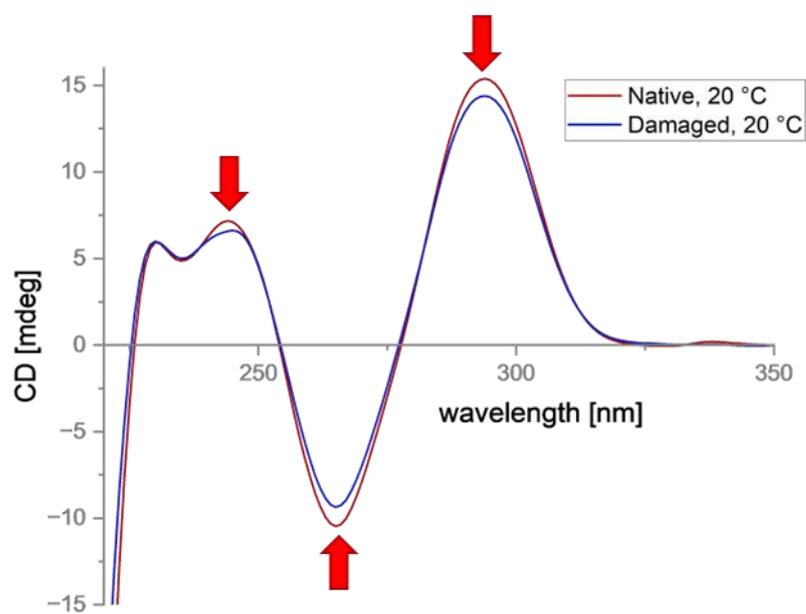


Figure A1: CD spectra of the G3T4 G-quadruplex before (red) and after (blue) the Fenton reagents (H_2O_2 1 mM, CuCl_2 200 μM , citric acid 1 mM) were added to the sample. The CD signal was measured at the start of the experiment and then two hours after adding the Fenton reagents. A decrease in the ellipticity at 240, 265 and 290 nm and were observed, which could be an indication that the G-quadruplex is destabilised and is unfolding as the reaction occurs.