Mechanisms and Dynamics of Oxidative DNA Damage Repair in Nucleosomes

Wendy J. Cannan
University of Vermont

Follow this and additional works at: https://scholarworks.uvm.edu/graddis

Part of the Biochemistry Commons, Genetics and Genomics Commons, and the Molecular Biology Commons

Recommended Citation
https://scholarworks.uvm.edu/graddis/628

This Dissertation is brought to you for free and open access by the Dissertations and Theses at ScholarWorks @ UVM. It has been accepted for inclusion in Graduate College Dissertations and Theses by an authorized administrator of ScholarWorks @ UVM. For more information, please contact donna.omalley@uvm.edu.
MECHANISMS AND DYNAMICS OF OXIDATIVE DNA DAMAGE REPAIR IN NUCLEOSOMES

A Dissertation Presented

by

Wendy Cannan

to

The Faculty of the Graduate College

of

The University of Vermont

In Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy
Specializing in Microbiology and Molecular Genetics

October, 2016

Defense Date: June 22, 2016
Dissertation Examination Committee:

David S. Pederson, Ph.D., Advisor
Stephen Everse, Ph.D., Chairperson
Sylvie Doublié, Ph.D.
Scott Morrical, Ph.D.
Susan Wallace, Ph.D.
Cynthia J. Forehand, Ph.D., Dean of the Graduate College
ABSTRACT

DNA provides the blueprint for cell function and growth, as well as ensuring continuity from one cell generation to the next. In order to compact, protect, and regulate this vital information, DNA is packaged by histone proteins into nucleosomes, which are the fundamental subunits of chromatin. Reactive oxygen species, generated by both endogenous and exogenous agents, can react with DNA, altering base chemistry and generating DNA strand breaks. Left unrepaired, these oxidation products can result in mutations and/or cell death. The Base Excision Repair (BER) pathway exists to deal with damaged bases and single-stranded DNA breaks. However, the packaging of DNA into chromatin provides roadblocks to repair. Damaged DNA bases may be buried within nucleosomes, where they are inaccessible to repair enzymes and other DNA binding proteins. Previous in vitro studies by our lab have demonstrated that BER enzymes can function within this challenging environment, albeit in a reduced capacity.

Exposure to ionizing radiation often results in multiple, clustered oxidative lesions. Near-simultaneous BER of two lesions located on opposing strands within a single helical turn of DNA of one another creates multiple DNA single-strand break intermediates. This, in turn, may create a potentially lethal double-strand break (DSB) that can no longer be repaired by BER. To determine if chromatin offers protection from this phenomenon, we incubated DNA glycosylases with nucleosomes containing clustered damages in an attempt to generate DSBs. We discovered that nucleosomes offer substantial protection from inadvertent DSB formation. Steric hindrance by the histone core in the nucleosome was a major factor in restricting DSB formation. As well, lesions positioned very close to one another were refractory to processing, with one lesion blocking or disrupting access to the second site. The nucleosome itself appears to remain intact during DSB formation, and in some cases, no DNA is released from the histones. Taken together, these results suggest that in vivo, DSBs generated by BER occur primarily in regions of the genome associated with elevated rates of nucleosome turnover or remodeling, and in the short linker DNA segments that lie between adjacent nucleosomes.

DNA ligase IIIα (LigIIIα) catalyzes the final step in BER. In order to facilitate repair, DNA ligase must completely encircle the DNA helix. Thus, DNA ligase must at least transiently disrupt histone-DNA contacts. To determine how LigIIIα functions in nucleosomes, given this restraint, we incubated the enzyme with nick-containing nucleosomes. We found that a nick located further within the nucleosome was ligated at a lower rate than one located closer to the edge. This indicated that LigIIIα must wait for DNA to spontaneously, transiently untrap from the histone octamer to expose the nick for recognition. Remarkably, the disruption that must occur for ligation is both limited and transient: the nucleosome remains resistant to enzymatic digest before and during ligation, and reforms completely once LigIIIα dissociates.
CITATIONS

Material from this dissertation has been published in the following form:

# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>CITATIONS</td>
<td>ii</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>vii</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>viii</td>
</tr>
<tr>
<td>CHAPTER 1: INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>1.1. Chromatin</td>
<td>1</td>
</tr>
<tr>
<td>1.1.1. Nucleosome Structure</td>
<td>1</td>
</tr>
<tr>
<td>1.1.2. Nucleosome Positioning &amp; DNA Sequence</td>
<td>2</td>
</tr>
<tr>
<td>1.1.3. Linker Histone</td>
<td>6</td>
</tr>
<tr>
<td>1.1.4. The 30 nm Fiber &amp; Higher Order Chromatin Structure</td>
<td>8</td>
</tr>
<tr>
<td>1.1.5. Nucleosome Dynamics</td>
<td>12</td>
</tr>
<tr>
<td>1.1.6. Nucleosome Energetics</td>
<td>15</td>
</tr>
<tr>
<td>1.2. DNA Damage</td>
<td>17</td>
</tr>
<tr>
<td>1.2.1. Non-oxidative Endogenous/Spontaneous DNA Damage</td>
<td>17</td>
</tr>
<tr>
<td>1.2.2. Oxidative Endogenous DNA Damage</td>
<td>19</td>
</tr>
<tr>
<td>1.2.3. Oxidative Exogenous DNA Damage</td>
<td>21</td>
</tr>
<tr>
<td>1.2.4. Oxidative Damage within the Mitochondria</td>
<td>22</td>
</tr>
<tr>
<td>1.2.5. DNA Double-Strand Break Formation</td>
<td>24</td>
</tr>
<tr>
<td>1.3. DNA Repair</td>
<td>26</td>
</tr>
<tr>
<td>1.3.1. Mismatch Repair (MMR)</td>
<td>26</td>
</tr>
<tr>
<td>1.3.2. Nucleotide Excision Repair (NER)</td>
<td>27</td>
</tr>
<tr>
<td>1.3.3. DNA Double-Strand Break Repair (DBSR)</td>
<td>29</td>
</tr>
<tr>
<td>1.3.4. Overview of Base Excision Repair (BER)</td>
<td>33</td>
</tr>
<tr>
<td>1.3.5. Glycosylases</td>
<td>35</td>
</tr>
<tr>
<td>1.3.6. AP Endonuclease 1 (APE1)</td>
<td>45</td>
</tr>
<tr>
<td>1.3.7. DNA Polymerase β (pol β)</td>
<td>49</td>
</tr>
<tr>
<td>1.3.8. Ligase IIIα &amp; XRCC1 (LigIIIα-XRCC1)</td>
<td>50</td>
</tr>
</tbody>
</table>
1.3.9. BER Accessory & Interacting Proteins .......................................................... 54
1.3.10. Choice Between BER Subpathways ......................................................... 58

1.4. DNA Repair in the Context of Chromatin ...................................................... 61
  1.4.1. Systems Used to Study BER in Nucleosomes ........................................ 61
  1.4.2. Effects of Lesion Orientation and Translational Position ...................... 64
  1.4.3. Pol β and Other BER Enzymes on Nucleosomes .................................. 70
  1.4.4. Linker Studies ....................................................................................... 72
  1.4.5. Histone PTMs and Nucleosome BER .................................................... 74
  1.4.6. Chromatin Remodelers and Nucleosome BER ..................................... 76

1.5. Figures .......................................................................................................... 79

1.6. Tables ............................................................................................................ 91

CHAPTER 2: NUCLEOSOMES SUPPRESS THE FORMATION OF
DOUBLE-STRANDED DNA BREAKS DURING ATTEMPTED BASE
EXCISION REPAIR OF CLUSTERED OXIDATIVE DAMAGES ......................... 95

2.1 Abstract ......................................................................................................... 95

2.2 Introduction .................................................................................................... 96

2.3 Experimental Procedures ............................................................................. 98

2.4 Results .......................................................................................................... 103

2.5 Discussion .................................................................................................... 114

2.6 References .................................................................................................... 117

2.7 Footnotes ...................................................................................................... 121

2.8 Table and Figure Legends ........................................................................... 122

2.9 Tables and Figures ....................................................................................... 130
CHAPTER 3: HUMAN DNA LIGASE IIIα FORMS A REVERSIBLE, LIGATION-PERMISSIVE COMPLEX WITH DNA NICK-CONTAINING NUCLEOSOMES

3.1 Abstract .......................................................................................................................... 139
3.2 Introduction ....................................................................................................................... 140
3.3 Results ............................................................................................................................... 143
3.4 Discussion ......................................................................................................................... 152
3.5 Experimental Procedures ............................................................................................... 157
3.6 References ......................................................................................................................... 163
3.7 Footnotes ........................................................................................................................... 167
3.8 Figure Legends .................................................................................................................. 168
3.9 Figures ............................................................................................................................... 179

CHAPTER 4: CONCLUSIONS AND FUTURE DIRECTIONS ........................................ 191
4.1. Conclusions ...................................................................................................................... 191
4.2. Future Directions ............................................................................................................. 193
  4.2.1. Could the nucleosome allow for complete clustered lesion repair? ............... 193
  4.2.2. What happens after a BER-induced DSB occurs? ........................................ 194
  4.2.3. Where do BER-induced clustered damages occur within the genome? ........ 195
  4.2.4. Could single molecule studies show us how LigIIIα finds nicks in DNA? ... 196
  4.2.5. What happens when LigIIIα encounters a nucleosome? ................................. 198
  4.2.6. Does LigIIIα exert torsion on DNA? ................................................................. 199
  4.2.7. Does SNF2H/SMARCA5 influence nucleosome BER? ............................... 200
LIST OF TABLES

Table                                    Page

Table 1-1. Formation and deactivation of reactive oxygen species (ROS) .................... 91
Table 1-2. Mammalian DNA Glycosylases .......................................................................... 92
Table 1-3. Nucleosome unwrapping rates. ......................................................................... 94
Table 2-1. The nucleosome suppresses single-strand break formation
relative to naked DNA. ......................................................................................................... 130
Table 2-2. Predicted vs. observed double-strand break formation. ................................. 131
### LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 1-1. Nucleosome structure &amp; histone fold domains.</td>
<td>79</td>
</tr>
<tr>
<td>Figure 1-2. Differing models for the 30 nm fiber</td>
<td>80</td>
</tr>
<tr>
<td>Figure 1-3. DSB formation via replication fork collapse.</td>
<td>81</td>
</tr>
<tr>
<td>Figure 1-4. Various products of endogenous reactive oxygen species or UV radiation</td>
<td>82</td>
</tr>
<tr>
<td>Figure 1-5. MMR in eukaryotes</td>
<td>83</td>
</tr>
<tr>
<td>Figure 1-6. NER in eukaryotes</td>
<td>84</td>
</tr>
<tr>
<td>Figure 1-7. DSBR pathways</td>
<td>85</td>
</tr>
<tr>
<td>Figure 1-8. BER and subpathways</td>
<td>86</td>
</tr>
<tr>
<td>Figure 1-9. DNA glycosylase catalysis, mono- and bi-functional</td>
<td>87</td>
</tr>
<tr>
<td>Figure 1-10. AP endonuclease 1</td>
<td>88</td>
</tr>
<tr>
<td>Figure 1-11. Ligase mechanism of action</td>
<td>89</td>
</tr>
<tr>
<td>Figure 1-12. LigIIIα-XRCC1</td>
<td>90</td>
</tr>
<tr>
<td>Figure 2-1. Double-strand break formation during attempted base excision repair (BER)</td>
<td>132</td>
</tr>
<tr>
<td>Figure 2-2. Impact of the stagger between opposing lesions on double-strand break formation</td>
<td>133</td>
</tr>
<tr>
<td>Figure 2-3. hNTH1-mediated, double-strand break formation in nucleosomes is lyase-limited</td>
<td>134</td>
</tr>
<tr>
<td>Figure 2-4. DSB formation due to processing of clustered lesions in 601 nucleosomes</td>
<td>135</td>
</tr>
<tr>
<td>Figure 2-5. Evidence for coupled processing of clustered lesions in nucleosomes</td>
<td>136</td>
</tr>
<tr>
<td>Figure 2-6. The efficiency of BER-dependent formation of DSBs in nucleosomes varies substantially among different DNA glycosylases</td>
<td>137</td>
</tr>
<tr>
<td>Figure 2-7. Fate of nucleosomes during the processing of closely opposed oxidative lesions</td>
<td>138</td>
</tr>
</tbody>
</table>
Figure 3-1. Spontaneous partial unwrapping of DNA from the histone octamer enables DNA ligase IIIα-XRCC1 to bind to DNA nicks in nucleosomes.................................................................................................................. 179
Figure 3-2. Ligation by LigIIIα-XRCC1 does not irreversibly disrupt nick-containing nucleosomes. ..................................................................................................................... 180
Figure 3-3. Binding of LigIIIα-XRCC1 does not irreversibly disrupt nucleosomes........................................................................................................................................ 181
Figure 3-4. Ligation occurs within the ternary complex that forms upon addition of LigIIIα-XRCC1 to nick-containing nucleosomes. .................................................... 182
Figure 3-5. Ligation of nicks in nucleosomes by LigIIIα-XRCC1 is not enhanced by linker DNA. ............................................................................................................................. 183
Figure 3-6. Reduced ligation efficiency in highly stable 601 nucleosomes. .............. 184
Figure 3-7. Ligation of nicks in nucleosomes by ligase IIIα does not require XRCC1. .................................................................................................................................. 185
Figure 3-8. Discrete steps in the recognition and ligation of DNA nicks in nucleosomes, by LigIIIα-XRCC1........................................................................................................... 186
Supplemental Figure 3-1. Carrier DNA does not interfere with ligation assays. ......................................................................................................................................................... 187
Supplemental Figure 3-2. Interactions between LigIIIα-XRCC1 and undamaged nucleosomes. .............................................................................................................................. 188
Supplemental Figure 3-3. Preferential suppression of interactions between LigIIIα-XRCC1 and undamaged nucleosomes at elevated salt concentrations. ...................................................................................................................... 189
Supplemental Figure 3-4. Extensive interactions between LigIIIβ and DNA................................................................................................................................................................................. 190
Figure 4-1. BER-induced double-strand break formation in nucleosomes............. 202
Figure A-1. APLF does not impact LigIIIα-XRCC1 Activity........................................ 247
CHAPTER 1: INTRODUCTION

1.1. Chromatin

1.1.1. Nucleosome Structure

In eukaryotic cells, nuclear DNA is compacted into chromatin by histones and non-histone architectural proteins. The fundamental unit of chromatin is the nucleosome, which consists of 145 to 147 bp of DNA wrapped ~1.7 times around an octamer of histone proteins in a left-handed manner (Luger et al., 1997). This “histone octamer” consists of two histone H3-H4 dimers which form a tetramer, with two histone H2A-H2B dimers flanking either end of the core, forming a sort of disk shape. The globular regions in the histones contain conserved “histone fold” domains, consisting of 3-α-helices separated by loops (Figure 1). Positively charged amino acids in the loop domains interact with the negatively charged sugar-phosphate backbone of DNA at each helical turn. Bending of DNA around the octamer decreases electrostatic repulsion on the outside of the bend, further stabilizing the structure. This wrapping serves the important purpose of compacting the DNA sequence ~5-fold. However, nucleosomes also create a regulatory barrier to DNA binding proteins and enzymes, as some of the DNA sequence is now “buried” towards the histone octamer and the entire sequence is now pinned down by histone-DNA contacts.
Histones also contain positively charged, disordered N-terminal “tails”. Approximately 25-30% of histone mass consists of disordered, N-terminal tails (H2A additionally has a disordered C-terminal tail), originally discovered due to their protease sensitivity (Bohm & Crane-Robinson, 1984). Histone H3 and H2B tails exit the nucleosome core between adjacent DNA superhelical gyres through “channels” formed by aligned, inward-facing minor grooves (Walker, 1984) (Luger et al., 1997). H2A and H4 tails, meanwhile, exit the nucleosome core over the top or bottom of the nucleosome, again following the track of minor grooves.

The histone tails are intrinsically disordered, containing numerous arginine and lysine residues (Hansen et al., 2006) which allow for additional contacts with nucleosomal DNA at physiological salt concentrations (Walker, 1984) (Smith & Rill, 1989). While removal of the tails generally increases accessibility for DNA binding proteins and enzymes, they contribute only modestly to overall nucleosome stability (reviewed in (Cutter & Hayes, 2015)). The histone tails are also subject to numerous posttranslational modifications, and are indispensable to higher-order chromatin structure, both discussed in additional detail later.

1.1.2. Nucleosome Positioning & DNA Sequence

The dyad axis marks the single base pair center of the DNA sequence wrapped around the octamer, near the interface of H3:H3. Here, the minor groove faces outward,
and is labeled superhelical position 0. Moving in either direction, each 10 bp helical turn is labeled as superhelical position +/-1, 2, 3, etc.. Thus, each nucleotide within the DNA sequence has its own, numbered position with respect to the dyad axis. Additionally, each nucleotide also has a “rotational” position, or how is oriented with respect to the histone octamer. We often refer to rotational position as “inward” or “outward” facing, wherein “inward” facing nucleotides are occluded by the close presence of the histone octamer, while “outward” facing nucleotides are located further from the histones. Rotational position has a periodicity of 10 bp, so if the minor groove faces outward nucleotide at superhelical position 0, it will also face outward at positions 1, 2, 3, etc. (see Figure 1 for a labeled view of the numbered superhelical locations). Given this ~10 bp periodicity, immediately adjacent nucleotides will differ in rotational position by ~36° (reviewed in (Cutter & Hayes, 2015)).

Proteins that recognize the major groove of DNA (such as restriction endonucleases) tend to be highly sequence specific, involving specific bonds between amino acids and bases. In contrast, the histone octamer interacts primarily with the DNA backbone, making it more sensitive to changes in groove geometry and bending rather than specific histone-DNA contacts (reviewed in (Widom, 2001)).

Different DNA sequences will have slightly different geometric properties. Some have an intrinsic bend in them, whereas others resist compression or bending, and some have wider or narrower minor grooves. In this respect, different DNA sequences can have
varying affinities for wrapping the histone octamer. A “5000-fold range of affinities” has been reported among a range of DNA sequences (Gencheva et al., 2006) (Thastrom, Bingham, et al., 2004) (Lowary & Widom, 1998). Drew & Travers examined intrinsic DNA bending as it relates to nucleosome formation by cutting a 169 bp length of circular DNA with the enzyme DNase I, whose cleavage capacity is sensitive to minor groove width (Drew & Travers, 1985). This revealed that minor groove sequences containing ATs tend to face inward (resisting cleavage by DNase I due to compression) and GC-containing minor grooves tend to face outward (resisting the compression necessitated by DNA circularization). Cleavage of a nucleosome containing this same sequence yielded a similar result, demonstrating that the same DNA compression is needed to efficiently wrap an octamer. They similarly probed genomic chicken DNA with DNase I, discovering that the genomic sequences were indeed, non-random. Again, they observed alternating runs of ATs (facing in) with stretches of GC (facing out), with a periodicity of about 10 bp, or one helical turn. Hydroxyl radical footprinting determined that the nucleosome decreases the average number of bases per helical turn in B-form DNA from 10.5 to 10.2 bp/turn (Hayes et al., 1990), which is consistent with later crystal structures (Luger et al., 1997).

DNA sequences that have intrinsic bending will be energetically favored in nucleosome formation, examples of which include: AA/TT/AT & CT dinucleotides spaced at 10 bp intervals; the GGGCCC motif, and the trinucleotide CTG repeat (Widom, 2001). Conversely, some sequences are refractory to DNA bending; for example, short, 5
base pair sequences and long stretches (10-20 bp) of poly(dA:dT) tracts are more likely to be excluded from nucleosomal sequences (Shrader & Crothers, 1989) (Field et al., 2008) (Yuan et al., 2005) (Lee et al., 2007) (Ozsolak et al., 2007) (Mavrich et al., 2008) (Kaplan et al., 2009).

Given the different affinities for nucleosome formation by different DNA substrates, it makes intuitive sense that DNA sequence might also influence nucleosome positioning within the entire genome. Also given the fact that nucleosome wrapping occludes many biological functions, as early as 1975 it was hypothesized that nucleosome position in the genome may correlate with expressed vs unexpressed genes, with unexpressed genes being more nucleosome-associated, and expressed genes having more free, non-nucleosome lengths of DNA (Oudet et al., 1975). In 1983, Pratt & Hattman performed DNA hybridization experiments with Tetrahymena, comparing the entire genome vs. "core-protected" (nucleosome-associated) DNA. They found that sequences containing N6-methyladenine residues were largely absent from the core-protected fraction of DNA; this was early definitive evidence for genome-wide nucleosome positioning (Pratt & Hattman, 1983).

We now know that cells possess chromatin remodelers and other agents that help enzymes and regulatory factors that bind DNA in chromatin. Nevertheless, in vivo nucleosome positioning itself can affect activity. For example, promoters and transcription factor binding sites tend to be excluded from nucleosomes, both in vitro and
in vivo. Studies have demonstrated that nucleosomes are, in at least some cases, lost from transcriptionally active promoters (Boeger et al., 2003) and that promoters of expressed genes in general tend to have nucleosomes less frequently associated with them than typically unexpressed promoter regions (Ozsolak et al., 2007). However, in higher eukaryotes, genome-wide nucleosome positioning studies have revealed that only a fraction of the genome is associated with well-positioned (with a standard deviation of ~10-20 bp), specifically located nucleosomes (often located upstream of an expressed gene promoter), underscoring the flexibility of DNA regulation in vivo (Zhang et al., 2015) (Schones et al., 2008) (Valouev et al., 2011).

1.1.3. Linker Histone

A nucleosome consists of 145 to 147 bp of core DNA directly contacting the histone octamer. However, a micrococcal nuclease digest of whole nuclei reveals that not all DNA is nucleosome associated (Noll, 1974). This non-octamer associated DNA, located in between adjacent nucleosomes, is referred to as “linker” DNA. The average linker length in the human genome is about 50 bp, making a single nucleosome (nucleosome DNA + linker DNA) ~200 bp long. However, linker lengths can vary from ~20 to 90 bp based on various factors, such as organism, cell type, and even within the same genome (Van Holde, 1988). For example, in rats, nucleosomes associated with telomere-proximal DNA have a shorter, 20 bp spacing (Makarov et al., 1993).
At least some linker DNA associates with its own histone(s), H1 (and H5 in the erythrocytes of birds and fish). Simpson first dubbed the term “chromatosome” to refer to a chromatin unit consisting of one histone octamer, one linker histone, and 160-170 bp of DNA; the result of a limited micrococcal nuclease digest where H1 has not been removed by high salt, although high micrococcal nuclease concentrations and/or longer digestion times will eventually cleave to the 147 bp core (Simpson, 1978). Linker histones have a lower affinity for DNA than the histone octamer; they can be removed by treatment with ~350 mM sodium chloride without removal of any octamer histones.

Histone H1 is larger (~20,000 Da) and less conserved than the core histones. However, although larger, only the ~80 residue central globular domain is well folded. H1 also contains a short unstructured N-terminus and a long (~100 residue) unstructured C-terminal tail. DNase I footprinting experiments found that linker histone protects anywhere from 15-30 bp of additional DNA, and this protection occurs at the dyad axis as well as entry/exit point of the nucleosome (Staynov & Crane-Robinson, 1988). This protection is usually symmetric, although a few studies found asymmetric protection depending upon the DNA sequence used. This led to a general model where histone H1 binds the minor groove of DNA somewhere near the dyad. Surprisingly, the disordered tails are largely dispensable for DNA protection of the entry/exit points from nucleases, however, the longer C-terminal tail of H1 is required for “stem” formation between nucleosomes. Under cryo-EM, this resembles the pinching together of DNA near the nucleosome near the entry/exit points (reviewed in (Crane-Robinson, 1997)). However,
the exact placement of linker histone with regard to the nucleosome has been a point of contention for over 30 years. Even today, there are conflicting models placing the globular domain either at the center of the nucleosome dyad (symmetric) or off-dyad (asymmetric). Recent cryo-EM and NMR studies place drosophila and human H1 linker histones off the dyad (Song et al., 2014) (Zhou, Feng, et al., 2013), while another study solving the crystal structure of the nucleosome bound to the globular domain of histone H5 places it at the center of the dyad (Zhou, Jiang, et al., 2015). Now, it seems clear that both “binding modes” can occur, and the latter study goes on to propose different consequences for higher order chromatin condensation based upon linker histone binding mode, with on-dyad binding compacting into a more condensed form of chromatin than off-dyad binding.

1.1.4. The 30 nm Fiber & Higher Order Chromatin Structure

The average human cell contains ~3 billion bp of DNA, which translates to ~2 meters if it were to be stretched out linearly. Fitting the genome into a ~10 µm nucleus is an astounding feat of packaging, requiring between 100 & 10,000 fold compaction. During mitosis, DNA is condensed even further, requiring over 20,000-fold compaction to achieve the structure we call the chromosome.

There are many players involved in the process of chromatin condensation and decondensation, including: architectural non-histone chromatin associated proteins,
nucleosome binding proteins, ATP-dependent chromatin remodelers, and histone chaperones. However, it’s worth noting that significant compaction can be achieved simply by the intrinsic properties of the histone octamer and, in particular, its histone tails. Internucleosomal interactions mediated by the histone tails are necessary for secondary (30 nm fiber) and tertiary (higher-order folding) chromatin condensation (Allan et al., 1982) (Schwarz & Hansen, 1994) (Garcia-Ramirez et al., 1995). Histone H4 tails interact with the H2A/H2B “acidic patch” on adjacent nucleosomes, contributing to condensed chromatin stability (Wilkins et al., 2014). Additionally, H3 contacts DNA on adjacent nucleosomes in condensed DNA (Zheng et al., 2005).

The condensation of the 10 nm fiber into the 30 nm fiber can be largely achieved solely through manipulation of mono- and divalent ion concentrations. At low salt concentrations (≤10 mM), cryo-EM of chromatin is represented by the 10 nm filament. The classical, but somewhat disorganized “beads on a string” appearance is due to the lack of linker histone. When histone H1 is present, the 10 nm fiber yields more of a zig-zag appearance, with H1 aligning the entry/exit points of DNA on the same side of each nucleosome. Increasing salt concentration condenses DNA, resulting in a 30 nm fiber with a coiled appearance at ~60 nM monovalent salt. H1 is unnecessary for condensation at higher salt, however; the fiber does not reach the same level of compaction without it (Thoma et al., 1979).
The structure of this highly condensed, 30 nm diameter fiber has been intensively studied and hotly debated over the years. Initially, the organization of the fiber was initially proposed as a coiled, “one-start” solenoid by the Klug group based upon EM (Figure 2). In the one-start model, each nucleosome interacts with the next nucleosome in sequence, compacting into a single helix. Additionally, due to this helical periodicity, interactions with every 5\textsuperscript{th} or 6\textsuperscript{th} nucleosome may be possible too. For this conformation, linker DNA would need to be bent. Williams et al. instead proposed a zig-zag, 2-start model based upon electron image and x-ray scattering data. In this model, alternating nucleosomes interact with each other, with linker DNA running relatively straight down the center (Williams et al., 1986). Unfortunately, EM measurements do not have the resolution needed to visualize the path of DNA in such a compacted structure, thus different methods would need to be employed to address the question.

In 2004, Dorigo et al. approached the problem by crosslinking arrays of 10-12 nucleosomes. Amino acids 14-19 of the histone H4 tail are necessary for contacting neighboring H2A/H2B molecules in the 30 nm fiber. By engineering cysteine H4 mutants in this regions, disulfide bonds could be introduced to specifically crosslink the secondary structure. Following compaction with linker histone H1 or MgCl\textsubscript{2}, the linkers were cleaved with restriction endonuclease Sca I. What was left was a fiber associated only by its crosslinks between nucleosomes. Following electrophoresis, they discovered that the 10-12 nucleosome arrays were cleaved into two 5-6 nucleosome pieces, concurrent with a two-start structure. Conversely, a single, 10-12 nucleosome piece
would have indicated a one-start fiber (Dorigo et al., 2004). One year later, the same lab crystallized a tetranucleosome at 9Å in the form of a two-start helix with zig-zag linkers (Schalch et al., 2005).

However, despite this strong evidence for the two-start model, the debate continues. Beyond the ever-present question of how well these highly-controlled experiments mimic \textit{in vivo} conditions, there are a number of factors that can affect the secondary structure of the 30 nm fiber. These include linker length, presence of H1 linker histone, and mono- and divalent salt concentrations. Consequently, more recent studies are now incorporating innovative modeling and simulation approaches to deal with the complexity of combining these multiple variables. Grigoryev et al. combined \textit{in silico} Monte Carlo simulations with \textit{in vitro} transmission EM of nucleosome arrays that has been covalently bound via limiting formaldehyde crosslinking to generate surprisingly similar results (Grigoryev et al., 2009). By both approaches, the result was a heteromorphic fiber that mainly contained a two-start zigzag structure, but was also interspersed with one-start conformations. Further “mesoscopic modelling” studies using simulation such as the one above have come to similar conclusions when taking in account multiple variables such as those listed earlier. It may be the case that the two-start model is favored either when linker DNA is short, or when linker H1 is present, both situations where the linker would remain relatively straight. The solenoid model, then, might be favored in situations where H1 is not present, or when linker DNA is long enough (≥70 bp), allowing for a bent conformation (reviewed in (Luger et al., 2012)).
As far as higher order chromatin structures go, we know that they exist – this is evident by the mitotic chromosomes, which have reproducible lengths, diameters, and banding patterns. However, there is no consensus as to how chromatin is organized within these structures. Research is limited in part due to resolution limits of both EM and light microscopy. It may be the case that the mitotic chromosome does not have a visibly ordered structure at all. Cryo-EM studies by Eltsov et al. performed on very thin slices of metaphase HeLa S3 cells yielded images of grainy and homogenous chromatin, with no indication of an ordered 30 nm fiber located within (Eltsov et al., 2008). They described this structure as a “polymer melt” or “molten core”. It may be that instead of a 30 nm fiber, highly condensed chromatin instead consists of “interdigitated” chains of nucleosomes that lose any kind of striking appearance once they are tightly packed together (reviewed in (Woodcock & Ghosh, 2010)).

1.1.5. Nucleosome Dynamics

The process of wrapping DNA around a nucleosome core creates steric hindrances to the DNA binding proteins and enzymes which act upon DNA. As has been noted previously, certain bases become “buried” away from sight, facing towards the histone octamer. There exist multiple proposed mechanisms by which a protein may overcome (or altogether avoid) this hindrance, including but not limited to 1) coupling of the protein to replication or transcription, 2) protein action only on exposed DNA sequences, and 3) active disruption/destruction of DNA-histone contacts. The Widom
group proposed a mechanism wherein spontaneous, partial unwrapping of DNA from the histone core would allow transient exposure of the target base/sequence. They initially tested this hypothesis through probing of nucleosomes with restriction enzymes targeted to specific sequences engineered within a nucleosome positioning sequence (Polach & Widom, 1995). Under this model, if the nucleosome were to undergo transient unwrapping, sequences positioned farther from the dyad axis would become transiently exposed more often than those sequences located closer to the dyad. This is what they observed as they calculated site exposure equilibrium constants further and further towards the dyad. They found that site exposure is reduced $10^2$ to $10^5$-fold as one moves from DNA at the edge of the nucleosome to the dyad. In a later study from the Widom lab, their hypothesis was refined to include the impact of DNA sequence: here, they took a DNA sequence that formed a more thermostable nucleosome (a synthetic, high-affinity nucleosome positioning sequence developed by the same lab, referred to simply as “601”) and calculated lower spontaneous unwrapping rates (Anderson & Widom, 2000). However, the fact that DNA at the dyad becomes exposed at all (albeit at a drastically reduced rate) suggests that all nucleosomal DNA – to some extent – is at least transiently available to the proteins which act upon it.

More recent studies examining nucleosome unwrapping rates have employed FRET (fluorescence resonance energy transfer) assays, wherein a donor dye molecule transfers its excitation energy to an acceptor molecule when the two molecules lie within ~1 to 10 nm of one another. The increase/decrease in acceptor dye emission can then be
correlated to the distance between dyes. Li & Widom engineered a dye pair between a Cy3 DNA (located at the edge of the nucleosome) and a nearby Cy5 histone acceptor (histones H3 and H2A were each tested). In the wrapped state, FRET efficiency is high, as distance between the two dyes is low. As a positive control, addition of sodium chloride lowered the FRET signal, indicating increased unwrapping rates. Nucleosomal DNA contained a binding site for E. coli repressor protein, LexA located a short distance into the nucleosome, and positioned as “inward” facing. Titration of LexA into the reaction resulted in a concurrent drop in FRET signal, providing further evidence of nucleosome unwrapping (Li & Widom, 2004). This is an example of stopped-flow FRET, which was used here to directly measure unwrapping rates. Rewrapping rates from this method may only be calculated indirectly, as binding of LexA precludes nucleosomes rewrapping. A subsequent study built upon this observation, incorporating the aforementioned stopped-flow FRET with fluorescence correlation spectroscopy (FCS). FCS measures FRET without the addition of a DNA binding protein; therefore, rewrapping rates dominate this reaction and are measured directly (Li et al., 2005). Together, the two methods yield both nucleosome unwrapping and rewrapping rates. From this, the Widom group concluded that, for a LexA binding site located 8–27 bp into the nucleosome, the wrapped nucleosome has a lifetime of only ~250 ms, with unwrapping events occurring every ~10-50 ms. A separate group reported faster unwrapping rates utilizing FRET (Tomschik et al., 2005). However, a follow-up experiment used crosslinked nucleosomes as a control, in which unwrapping cannot occur. This revealed that these apparently faster unwrapping rates were the result of
arteфactual “photoblinking”, a phenomenon wherein the acceptor dye intermittently stops emitting on a μs to ms time scale (which unfortunately, was exactly the timescale being observed) (Tomschik et al., 2009).

Further studies by the Widom lab examined additional translational positions located farther in the nucleosome. Moving the binding site 10 bp further in (18-37 bp) increased wrapped nucleosome lifetime from 250 ms to ~1 min, while a binding site located 20 bp further (28-47 bp) only increased wrapped lifetime another 10-fold, to ~10 min (Tims et al., 2011) (results are summarized in Table 3).

1.1.6. Nucleosome Energetics

There are many proteins that can exhibit force/torque on nucleosomes, and are not limited to chromatin remodelers. Helicases unwind DNA with a force of >20 pN (Bianco et al., 2001), while *E. coli* RNA polymerase exerts up to 40 pN (Wang et al., 1998). Microtubule- and actin-filament-associated molecular motors typically generate <10 pN of force. RSC and SWI/SNF have been measured on nucleosome arrays to exert ~12 pN of force (Zhang et al., 2006). But how much energy does it take to remove DNA? And is the energy required to unwrap DNA from the nucleosome the same, regardless of DNA sequence or location?

These questions have been addressed in single molecule studies involving molecular tweezers and optical traps. The general experimental setup involves a length
of DNA containing one or more nucleosomes, with one end bound to the source of force (often a cover slip, or a bead attached to a micropipette), and the other end bound to a bead in an “optical trap”. The optical trap is a laser which measures the distance the bead moves in response to force, and from this indirect measurement, force can be calculated. Brower-Toland et al. examined 5S rDNA nucleosome arrays, observing that DNA was released in multiple stages. The 1st stage represents the release of 76 bp of “outer wrap” DNA, and required ~5 pN to achieve. The 2nd stage required a higher force (~15 pN), and represented unwrapping of the 80 bp “inner wrap” of DNA, and itself was broken down into steps: the DNA is unwrapped first with the histones still attached, followed by release of the histones and complete nucleosome destruction (Brower-Toland et al., 2002). A subsequent, similar study (Mihardja et al., 2006) used mononucleosomes and 601 DNA, but found similar forces required: 3 pN for the outer wrap and 8-9 pN for the inner wrap. They calculated ΔG unwrapping of the outer DNA to be ~22 kJ/mol. Both of these studies also demonstrated that under low force, the nucleosome can reform after DNA is forcibly unwrapped. Mack et al. also employed the molecular tweezers method to compare normal nucleosomes to those containing histone “sin” mutants that reduce nucleosome stability (in this case, an H4 R45H mutation). (A sin nucleosome contains a mutant histone which has been demonstrated to allow transcription start without the need for SWI/SNF remodeling (Kruger et al., 1995)). Here, they not only observe that less force is required to unwrap a sin mutant histone, but also demonstrated that inner wrap unwrapping is reversible as well, so long as the histones are not displaced. From this
unwrapping/rewrapping equilibrium, they could calculate the absolute free energy for the inner wrap to be ~80 kJ/mol for WT nucleosomes (Mack et al., 2012).

1.2. DNA Damage

1.2.1. Non-oxidative Endogenous/Spontaneous DNA Damage

Even in the absence of external damaging agents, the physical integrity of DNA is constantly being threatened by its cellular environment. It is estimated that the nitrogenous bases that make up our DNA are subject to ~10,000 lesions per day per cell (reviewed by recent Nobel Laureate Tomas Lindahl in (Lindahl, 1993). While water molecules are obviously essential to the cell (and life, for that matter), they also catalyze multiple forms of DNA damage. The N-glycosyl bond between a base and sugar molecule is particularly sensitive to hydrolysis, resulting in spontaneous base loss estimated at 2000-10,000 per cell generation. Purines are far more susceptible to hydrolysis than pyrimidines (which is why one hears the term “depurination” far more often than “depyrimidination”). Left unrepaired, abasic sites (also referred to as apurinic/apyrimidinic or AP sites) can stall replicative polymerases, ultimately generating a double-strand break (DSB) if the stalled fork is not resolved (see Figure 3) (Higuchi et al., 2003). In some cases, the AP site can be bypassed by translesion synthesis, however, an adenine is preferentially placed across from the site, likely creating a mutation (Shibutani et al., 1997).
Conversely, the pyrimidine cytosine and its common methylated form 5-methylcytosine are subject to deamination through hydrolysis. Deamination of cytosine results in a uracil, while 5-methylcytosine is deaminated to thymine. Left unrepaired, these deamination products will produce a G:C to A:T transition mutations upon replication. However, 5-methylcytosine deamination occurs at a higher rate, and is made even more deleterious due its slow repair, as it must be processed by the mismatch repair pathway (which is typically slower than base excision repair) (reviewed in (Lindahl, 1993)). Consequently, not only are methylated regions of the genome are more prone to mutations, but CpG islands have become lost over time; CpG dinucleotides currently represent ~1% of the human genome, whereas random probability would predict an occurrence of ~4% (Lander et al., 2001).

Erroneous, spontaneous methylation of bases can occur through endogenous alkylating agents. The small molecule S-adenosylmethionine (SAM), serving as a methyl donor in enzymatic methylation of DNA, can also non-enzymatically react with DNA (Rydberg & Lindahl, 1982). The major alkylation product of SAM is 7-methylguanine, estimated at ~4000 alkylations per day per mammalian cell, though it can also result in 3-methyladenine and O6-methylguanine residues. The generation of 7-methylguanine in and of itself is not deleterious (it is not mutagenic or polymerase stalling). However, 7-methylguanine is more unstable than guanine and can generate AP sites or an open ring
structure, both of which can result in replication fork stalling (reviewed in (De Bont & van Larebeke, 2004)).

1.2.2. Oxidative Endogenous DNA Damage

Eukaryotic cells require oxygen to carry out many of their chemical reactions, not the least of which is energy production within the mitochondria. However, the reactivity of oxygen is a double-edged sword, making it the most frequent source of endogenous DNA damage. Reactive oxygen species (ROS), reviewed in (De Bont & van Larebeke, 2004), are particularly reactive molecules, and are generated as byproducts of multiple cellular pathways. ROS include hydrogen peroxide (H$_2$O$_2$), superoxide (O$^{-}$), singlet oxygen (¹O$_2$), and the hydroxyl radical (-OH) (Table 1). A major source of ROS in cells is the electron transport chain in mitochondria, which produces superoxide radicals. Nitric oxide (NO), which is used as a signaling molecule and, in phagocytes, helps to kill pathogens, is another endogenous source of oxidative damage. Reaction of NO with superoxide can result in formation of the ROS peroxynitrite (ONO$_2$^{-}).

The cell has mechanisms in place for dealing with endogenous ROS before they cause damage to the cell. Superoxide dismutases (SOD) convert superoxide to the less reactive hydrogen peroxide, which can then be converted to water and oxygen by peroxiredoxins and glutathione peroxidase; catalase can catalyze this reaction as well, although it is not found within the mitochondria (Table 1). However, left unprocessed, hydrogen peroxide can undergo a Fenton reaction with metal ions such as Fe$^{+2}$, which
generates the extremely reactive hydroxyl radical. Sequestering of DNA in the nucleus, away from the oxidative environment of the mitochondria, provides another level of defense, as the high reactivity of ROS also means they will not travel far before reacting with nearby molecules. For example, the half-life of a hydroxyl radical is about one nanosecond, yielding a reactive radius of ~20Å (reviewed in (Halliwell, 2001)). Tight packaging of DNA into nucleosomes and chromatin provide significant defense as well. Histones and other chromatin associated proteins act as a barrier to ROS by sterically occluding reactive DNA moieties. Additionally, proteins may react with or scavenge free radicals. Protection by chromatin has been demonstrated in multiple studies in which chromatin-associated proteins are removed in a step-wise manner with increasing salt (reviewed in (Cannan & Pederson, 2016)). Complete removal of protein yields up to a 70-fold increase in ROS-generated DNA damage. As well, chromatin condensation per se may offer an additional level of protection, possibly by increasing the local concentration of protein surrounding DNA (reviewed in (Lavelle & Foray, 2014)).

Despite multiple safeguards, oxidation of DNA bases occurs quite frequently. The major purine oxidation product in the cell is 8-hydroxyguanine (8-oxoG), and is often used as a marker for overall oxidative damage/stress (Melvin et al., 1998). Steady state levels have been estimated at one 8-oxoG per \(10^5-10^6\) DNA bases (Escodd, 2002). Predictably, 8-oxoG levels increase when the cell is subjected to oxidative stress (Malins & Gunselman, 1994). 8-oxoG is mutagenic because it can mispair with adenine during
replication, resulting in a G to T transversion (Minowa et al., 2000) (Klungland et al., 1999).

The major pyrimidine oxidation product is thymine glycol (Tg), resulting from the oxidation of thymine. This lesion, in contrast to 8-oxoG, can stall replicative polymerases, which can lead to a strand break during replication (Aller et al., 2007). It’s worth noting that the deamination of 5-methylcytosine and subsequent inefficient mismatch repair have commonly been viewed as the primary driver of the C to T transitions observed in CpG islands. However, oxidation of 5-methylcytosine creates a 5-methylcytosine glycol intermediate, which is then deaminated to thymine glycol, creating another pathway to mutation (Zuo et al., 1995).

Besides 8-oxoG and Tg, there many other varieties of oxidized bases. Figure 4 lists some of the many additional base oxidation products. (Bjelland & Seeberg, 2003) provide a more exhaustive list of known oxidative base damages. Beyond base oxidation, ROS can react with other portions of DNA, including the N-glycosidic bond (resulting in an abasic site) and the DNA backbone itself (resulting in a strand break).

1.2.3. Oxidative Exogenous DNA Damage

Although there are numerous physical and chemical insults that can damage the genome in a variety of ways, this chapter will focus primarily on exogenous oxidative damage. Chemical agents include the aforementioned H₂O₂ and NO as well as
hypochlorous acid (HOCl). Environmental agents that create chronic inflammation (such as smoking and *Helicobacter pylori* infection) can also lead to oxidative damage through neutrophil oxygen bursts of O$_2^-$ and H$_2$O$_2$. However, the most extensively studied exogenous oxidizer is ionizing radiation (IR), common sources of which include: X-Rays, gamma rays, cosmic rays, and radon emissions. IR is defined as any subatomic particle or electromagnetic wave with sufficient energy to remove electrons from atoms.

Oxidative damage from IR can occur either directly or indirectly, but mostly indirectly through its ability to separate water molecules into H$^+$ and hydroxyl radicals. This was determined by comparing the effect of IR on DNA in a dimethyl sulfoxide (DMSO) solution vs. DNA in water, with the DNA/DMSO resulting in far fewer damages than DNA/water (Weinfeld & Soderlind, 1991). DNA damage resulting from IR has some unique properties compared to DNA damage resulting from most chemical oxidizers, and tend to result in particularly deleterious “clustered damages” which will be further discussed in sections dealing with DNA double-strand breaks as well as in Chapter 2.

**1.2.4. Oxidative Damage within the Mitochondria**

It may seem a foregone conclusion that mitochondrial DNA (mtDNA), which lacks the protection of histones, is more susceptible to ROS due to its proximity to the electron transport chain. Indeed, the Free Radical/Mitochondrial Theory of Aging, proposed 60 years ago, and refined over the years, postulates that accumulated ROS
damage to the mitochondrial genome is the major cause of aging ((Harman, 1956) and reviewed in (Alexeyev, 2009)). Although mitochondrial DNA accumulates mutations ~10-fold faster than genomic DNA, these could be due to other factors, such as the faster doubling rate of mitochondria or the relatively poor fidelity of the mitochondrial replicative polymerase (reviewed in (Shokolenko et al., 2014).

A specific problem in determining the role of ROS in mitochondrial aging is the difficulty in pinning down actual levels of oxidative damage in mtDNA. A 1988 study put 8-oxoG levels at ~15-fold higher in mtDNA than in the nucleus (Richter et al., 1988). However, the same group a decade later found that the initial study suffered from artefactual oxidation (Helbock et al., 1998). Furthermore, studies from other groups have found little to no difference in oxidation levels between mtDNA and genomic DNA (Anson et al., 1999) (Anson et al., 2000) (Richter et al., 1988). Although mitochondria lack histones, they contain high concentrations of other proteins, which could substitute for the protection offered by chromatin. Isolated mitochondrial nucleoids have been shown in vitro to protect mtDNA from H$_2$O$_2$ and X-Rays (Guliaeva et al., 2006). In particular, mitochondrial transcription factor A associates tightly with mtDNA, and is present in large enough amounts to completely cover the mitochondrial genome (Alam et al., 2003). The mitochondria lack a full complement of DNA repair machinery; in particular, there is no known nucleotide excision repair occurring in mtDNA. However, the nuclear genome does export a subset of repair enzymes. In particular, enough of the base excision repair pathway is transported to the mitochondria to complete base excision
repair (reviewed in (Prakash & Doublié, 2015)). In fact, at least one study observed more efficient repair of 8-oxoG within the mitochondria than in nuclear DNA (Thorslund et al., 2002). Taken together, these caveats are compelling evidence to perhaps refine the Mitochondrial Theory of Aging and the current mode of thinking.

1.2.5. DNA Double-Strand Break Formation

DNA double-strand breaks (DSBs) are considered the most deleterious form of DNA damage to the cell, as they can lead to genome fragmentation and cell death if left unrepaired (Brown & Baltimore, 2000). Endogenously generated DSBs can occur via replication stalling (see aforementioned fork-collapse), transcription blocks, topoisomerase II, and aborted repair intermediates. Non-deleterious, programmed DSBs occur as well, such as during meiosis I recombination and immunoglobulin class switching in B-lymphocytes (for reviews, see (Soulas-Spraul et al., 2007) (de Massy, 2013)). Exogenous DSB formation occurs largely through IR (reviewed in (Barnard et al., 2013)), but can also be generated by radiomimetic cancer drugs such as bleomycin and topoisomerase II inhibitors (Povirk, 1996) (Hande, 1998) (Chen & Stubbe, 2005) (Koster et al., 2007).

Unlike H\textsubscript{2}O\textsubscript{2}-generated ROS that tend to induce single, isolated lesions, IR passes through the nucleus in a straight track, transferring energy to nearby molecules (e.g. water) as it travels. In this manner, a single photon or \textalpha\textendash particle can generate multiple ROS along its path. As IR energy increases, ionizing events become more frequent and
consequently, ROS become more closely spaced. These will then react with DNA, inducing closely opposed lesions which are commonly referred to as “clustered damages” (Goodhead et al., 1993) (Rydberg, 1996) (Sutherland et al., 2000). These damages consist of two or more DNA damages located within one or two helical turns of DNA. DSBs can arise spontaneously from clustered lesions if, for example, two of these lesions are single-stranded breaks (SSBs) located on opposite DNA strands. Because the base excision repair (BER) pathway (see Section 1.3.4) also generates SSBs as intermediates, this means that base lesions can also be converted to DSBs via near-simultaneous, attempted BER. This phenomenon is reviewed in (Cannan & Pederson, 2016) and discussed in more detail in Chapter 2.

IR- and ROS-generated strand breaks (both single and double) are often “dirty”. That is, the DNA ends associated with these breaks do not have perfectly re-ligatable 3′OH and 5′P groups. Instead, they commonly contain a variety of blocking moieties that are refractory to repair (Weinfeld & Soderlind, 1991). ROS-induced breaks can yield: 3′-phosphate (3′-P), 3′-phosphophoglycolate (3′-PG), 5′ phosphate (5′-P), 5′-aldehyde and 5′-hydroxyl (5′-OH) moieties. No single enzyme can process all of these groups. Instead there are a variety of repair enzymes from multiple pathways involved in the processing of these so-called dirty breaks (reviewed in (Andres et al., 2015)). It is likely that damage recognition by these enzymes (as well, whether the break is a SSB, DSB, or complex, clustered damage) influences downstream repair pathway choice, discussed in later sections.
1.3. DNA Repair

My focus on DNA repair will center on the repair of oxidative damages via base excision repair and its subpathways. Additionally, DSB repair will be covered to some extent, as it is employed to repair DSBs created inadvertently by BER processing of clustered lesions. I will only briefly touch upon mismatch repair and nucleotide excision repair. As well, I will devote more attention to repair in eukaryotes than in prokaryotes such as *E. coli*, as they are more relevant to repair within the context of chromatin.

1.3.1. Mismatch Repair (MMR)

MMR exists to repair single base-base mismatches and short insertion/deletion loops in DNA. Single nucleotide mismatches can result from the aforementioned deamination of 5-methylcytosine, while insertion/deletion mismatches arise from polymerase slippage and recombination intermediates.

See (Jiricny, 2006) & (Modrich, 2016) for detailed overviews of the MMR process. Briefly, in humans, 1-2 nt mismatches are recognized by *E. coli* MutS homologs MSH2/MSH6, while loops are recognized MSH2/MSH3. The MutL homolog MLH1 and its binding partners, post-meiotic-segregation increased proteins 1 & 2 (PMS1 and PMS2) appear to coordinate repair; the complex they form possesses an endonuclease activity that initiates degradation of the mismatched strand by exonuclease I (and
possibly other exonucleases). Subsequent to this, replication machinery, including the PCNA clamp, the RFC clamp loader, DNA polymerase δ (pol δ) and DNA ligase I (Lig I) synthesize DNA to replace the degraded strand (Figure 5).

In prokaryotes, the newly synthesized daughter strand is recognized by MMR by its lack of methylation (Pukkila et al., 1983). In eukaryotes, it’s still unclear how MMR is directed to the nascent, mismatched strand rather than the template strand. The human MMR machinery recognizes DNA nicks and ends, which offers explanation of how repair could occur on the newly synthesized, unligated lagging strand (Holmes et al., 1990). It is less clear how this might occur on the leading strand, although nicks arising from events such as the restart of a stalled replication fork could provide a proper substrate for recognition (Heller & Marians, 2006).

1.3.2. Nucleotide Excision Repair (NER)

NER enzymes recognize and process a variety of “bulky” lesions in DNA (reviewed in (Friedberg, 2006)), including: the intrastrand crosslinks induced by UV-B radiation such as cyclobutane pyrimidine dimers (CPDs) and pyrimidine 6-4 pyrimidone photoproducts (6-4PP). NER enzymes can also recognize cisplatin-induced crosslinks and DNA adducts formed by genotoxic agents such as benzo[a]pyrene and acetylaminofluorene. In bacteria and most eukaryotes, but not in placental mammals, CPDs and 6-4PPs can be directly reversed using the energy of light photons by CPD photolyase and 6-4 photolyase, respectively, in a process known as photoreactivation.
Humans, however, must rely on NER to repair UV DNA crosslinking damage. NER deficiencies can result in xeroderma pigmentosum (XP), a disease which confers an extreme sensitivity to solar UV irradiation and a $>200$-fold increased risk of skin cancer (reviewed in (DiGiovanna & Kraemer, 2012)). There are two major forms of NER: global genome NER (GG-NER) vs. transcription-coupled NER (TC-NER). As the names imply, GG-NER can take place in all regions of the genome, while TC-NER repairs lesions found within actively transcribed genes. The two pathways have different initiating factors, but the core factors remain the same.

Refer to Figure 6 and (Scharer, 2013) for an overview of NER. Briefly, GG-NER is initiated by XPC-RAD23B, which recognizes distorted DNA rather than any specific lesion, making NER a particularly versatile repair pathway. However, because CPDs actually do not disrupt the DNA helix to any great extent, UV-damaged DNA-binding protein (UV-DDB), which specifically recognizes CPDs, is sometimes required in conjunction with XPC-RAD23B. TC-NER, in contrast is initiated when RNA polymerase stalls at the site of a bulky lesion, and involves the TC-NER specific factors CSA, CSB, and XAB2, which are specific to TC-NER. Following recognition, TFIIH, which is also a transcription factor, is recruited through direct interaction with XPC-RAD23B. TFIIH is a large, 10 subunit complex which contains both a core kinase and a cyclin-activated kinase (CAK) that is not required for NER. One helicase subunit of TFIIH, XPB, initially pries open the DNA; XPD then unwinds the DNA until it stalls at the bulky lesion. This unwinding and stalling triggers recruitment of a “pre-incision”
complex, consisting of XPA, RPA, and XPG, and the concurrent dissociation of XPC-RAD23B. XPA appears to act as the complex “coordinator”, as it interacts with both upstream and downstream components, and is highly regulated at both the transcriptional level and by post-translational modifications. ERCC1-XPF is subsequently recruited by XPA, and dual incision takes place on the DNA backbone, with ERCC1-XPF, initiating incision on the 5’ end, followed by 3’ cleavage by XPG. At this point the replication factors, pol δ/ε, the PCNA clamp, the clamp loader RFC, RPA, and DNA ligase I purportedly step in and take over repair, although surprisingly, translesion synthesis pol κ has been implicated in NER (Ogi & Lehmann, 2006). Interestingly, DNA ligase IIIα/XRCC1 appears to perform the final ligation step in quiescent cells, with DNA ligase I functioning in NER only in actively dividing cells (Moser et al., 2007).

1.3.3. DNA Double-Strand Break Repair (DBSR)

The challenges of DSB repair are two-fold. Firstly, the correct two DNA ends must be joined, as incorrect end-joining will result in chromosome translocations and potentially cell death. Secondly, DNA must be restored to its correct sequence; failure to do so will result in sequence mutations. Ionizing radiation in particular makes this second challenge particularly difficult, as IR-induced DSBs often result in base loss at the site of strand break (reviewed in (Jackson & Bartek, 2009)). This sections provides only brief overviews of the DSB repair pathways. Refer to (Iliakis et al., 2015) for an in-depth comparison of Homologous Recombination Repair vs. Non-Homologous End-Joining vs. Alternative Non-Homologous End-Joining.
Homologous recombination repair (HRR) begins with 5’ to 3’ end resection, and involves numerous factors, including MRN, CtfP, Exo1, Dna2 and BLM helicase (Figure 7 and reviewed in (Symington & Gautier, 2011)). End resection represents a point of no return: NHEJ can no longer occur once this step occurs. The resulting 3’ ssDNA is bound by replication protein A (RPA), protecting the structure from nucleases. Rad51 forms a filament on the 3’ ends, and requires various accessory proteins to do so, including Rad51B, Rad51C, Rad51D, XRCC2, XRCC3, and BRCA2 (San Filippo et al., 2008). During synapsis, when two homologous chromosomes are matched up, Rad51 performs a search for homology on the sister chromatid. Once found, Rad51 initiates strand invasion, then dissociates with the help of Rad54. Some subpathways of HRR create a Holliday junction that must be resolved with resolvases and Lig I.

The major advantage of the HRR pathway over Non-Homologous End-Joining is that it is error free when carried out properly (Moynahan & Jasin, 2010). However, repair requires an undamaged template. In diploid organisms, HRR could theoretically utilize either the homologous chromosome or the replicated sister chromatid as the template for repair. But in practice, recombination between homologs in humans only occurs during meiosis I in germ lines. The sister chromatid is the preferred substrate for HR due to speed and efficiency; it is already proximal and aligned with the damaged chromosome following replication. However, this also means that HRR is restricted to S and G2 phases of the cell cycle, when the sister chromatids are present and not segregated to
opposite ends of the cell (Rothkamm et al., 2003) (Mladenov et al., 2013). In humans, few cells are actively dividing. Even most stem cells are not actively proliferating until stimulated to do so. Therefore, HRR is largely not a viable option. This also means that detecting HRR defects in higher eukaryotic cell lines can be challenging due to its limited role. Conversely, HRR is the predominant pathway in simple, single-celled eukaryotes such as yeast (Resnick & Martin, 1976) (Lewis et al., 1999).

If the advantage of HRR is its error-free nature, then its main disadvantage is repair time. In yeast, where HRR predominates, DSBR in yeast is extremely slow, requiring hours and several generation times (Frankenberg-Schwager et al., 1980). Conversely, the major advantage to Non-Homologous End-Joining (NHEJ), described below, is its speed. Mammalian cells repair IR-induced DSBs with half-times of 15-30 min, reflective of the major pathway in higher eukaryotes, NHEJ (Dahm-Daphi & Dikomey, 1996) (Schipler & Iliakis, 2013).

**Classical Non-Homologous End-Joining (c-NHEJ or NHEJ).** In eukaryotic NHEJ, broken DNA ends are recognized by the Ku70/80 heterodimer (Figure 7 and reviewed in (Symington & Gautier, 2011)). Recruited DNA-PKcs dimerizes at the site of the break, acting as a scaffold for subsequent proteins. As described in the section of DSB formation, IR tends to create “dirty breaks” with ends that are not initially suitable for ligation; when this is the case, the DNA ends must be processed and/or filled in by various accessory factors, including the bypass DNA polymerases μ/λ, polynucleotide
kinase and the Artemis nuclease (Lieber, 2010a). One the ends are sufficiently processed, Ligase IV (complexed with XRCC4) ligates the ends together. Additional factors in NHEJ include: XLF/Cernunnos, which interacts with XRCC4 and stimulates the reaction as well as PAXX, which stabilizes the complex and interacts with Ku (reviewed in (Iliakis et al., 2015)). Yeast and bacteria also have different, respective forms of NHEJ, but they are not essential to cell survival; the NHEJ pathway can be knocked out of yeast without generating a defective repair phenotype (Dudasova et al., 2004).

As NHEJ is template-independent, it is more error prone than HR, particularly with regard to IR-induced breaks. Even restriction endonuclease-generated breaks (which are clean and readily ligatable ends) are not always repaired in an error-free fashion through NHEJ (Lieber, 2010b). Additionally, there does not appear to be any mechanism in place to ensure that the correct two ends are joined by NHEJ, other than proximity. Therefore, chromosome translocations are more likely to occur with NHEJ than with HR. However, translocations are still less likely to occur in c-NHEJ than in alt-NHEJ (discussed below).

**Alternative Non-Homologous End-Joining (alt-NHEJ).** Inhibition of c-NHEJ in mammals reveals a DSBR activity that is qualitatively different from HRR. Dubbed alt-NHEJ (Figure 7), this form of DSBR repair is slower, error-prone, and results in chromosome translocations (reviewed in (Iliakis et al., 2015)). In addition to DSBR, alt-NHEJ has been implicated in class switch recombination (CSR) and, under certain
conditions, V(D)J recombination (Yan et al., 2007) (Corneo et al., 2007). A host of proteins have been associated with alt-NHEJ function, most of which already have identified repair capacities within the cell. PARP-1, which recognizes single-strand DNA breaks, and the helicase WRN bind the break first, followed by end resection by CtIP and Mre11. Pol θ promotes end-joining and microhomology annealing (Kent et al., 2015), and LigIIIα-XRCC1 (or LigI) completes repair (Wang et al., 2005) (Liang et al., 2008). Linker histone H1 acts as a stimulatory element to the process by enhancing Ligase IIIβ end-joining in vitro (Rosidi et al., 2008). The exact mechanism behind this stimulation is unknown, but it is proposed that H1 acts to bind and align the DNA ends for ligation. A recent study screening a host of DNA damage response proteins with siRNA found a host of knockdowns that resulted in reduced alt-NHEJ activity yielded an additional 13 genes potentially involved in alt-NHEJ (Howard et al., 2015). Because so many proteins are implicated in alt-NHEJ, and because they all seem to have other functions, it may be that alt-NHEJ represents more of a last ditch effort than a single, organized pathway. It’s possible that the particular players involved in alt-NHEJ vary based on factors such as break location (e.g. euchromatin will contain a different set of proteins vs. heterochromatin). It may appear striking that so many mechanisms exist in the mammalian cell to suppress DSB persistence. However, complete failure to complete repair would lead to genomic fragmentation and ultimately cell death.

1.3.4. Overview of Base Excision Repair (BER)
Nucleotide base damages, as well as single-strand breaks and abasic sites, are all funneled into the BER pathway. Left unrepaired, these damages can lead to mutations, replication-associated double-strand breaks, and/or cell death. In humans, deficiency of the BER machinery has been linked to multiple human diseases, including premature aging, neurodegeneration, metabolic syndromes, and cancer (for reviews, see (Vijg & Suh, 2013) (Canugovi et al., 2013) (Wallace et al., 2012))

The initial step in BER of damaged bases is recognition of the base by its corresponding glycosylase, followed by DNA binding and nucleotide excision through cleavage of the N-glycosylic bond (Figure 8 and reviewed in (Wallace, 2014)). Bifunctional glycosylases have an additional AP lyase activity that cleaves the DNA backbone, whereas monofunctional glycosylases leave the backbone intact. Following base removal, in humans, AP-endonuclease 1 (APE1) cleaves the DNA backbone 5’ to the abasic site. If the backbone has already been cleaved by a glycosylase, the DNA ends are processed by either APE1 or polynucleotide kinase (PNK), resulting in a single nucleotide gap. In short-patch repair, this gap is filled by DNA polymerase β (pol β), and the remaining nick in the DNA backbone is sealed by DNA Ligase IIIα and its scaffolding protein partner XRCC1 (LigIIIα-XRCC1). Alternatively, in long patch repair, pol β, pol δ, and/or pol ε continue to incorporate 2-6 nucleotides onto the 3’ end of the repaired strand, displacing the downstream parent strand into an unannealed “flap” that requires removal by flap endonuclease 1 (FEN1). DNA ligase I then repairs the resulting nick.
The following sections will explore the individual steps and proteins of BER in further detail, and will also touch upon BER subpathways, including accessory proteins and interaction partners, and the choice between repair subpathways. In the final chapter, I will examine BER in the context of nucleosomes, and the challenges presented by chromatin structure.

1.3.5. Glycosylases

The first step in BER is excision of the damaged base by a glycosylase. Since a wide variety of base damages can result from varying types of damage and differing base chemistries, there exists a fairly extensive selection of DNA glycosylases – over 10 in humans - that excise specific damaged bases (summarized in Table 2). In contrast, subsequent BER steps typically require only one specific enzyme (with few or no alternatives) acting upon a BER intermediate.

DNA glycosylases are relatively small, ranging from ~15 to 60 kDa. Upon damage recognition, the glycosylase binds the minor groove of DNA, extruding or “flipping” the damaged base out of the DNA helix via the major groove into an active site (reviewed in (Brooks et al., 2013)). This base-flipping action is required for lesion recognition by nearly all known glycosylases, although recently, an exception to this rule has been discovered with the crystal structure of the glycosylase AlkD (Mullins et al., 2015). The subsequent removal of the base lesion results in an abasic or...
apurinic/apyrimidinic (AP) site. While many glycosylases are monofunctional, leaving an abasic site, the glycosylases that remove oxidative damages are bifunctional, and can subsequently cleave the DNA backbone following base excision, although with varying degrees of efficiency (reviewed in {Kim, 2012 #498}).

The glycosidic mechanism of action is differs between mono- and bifunctional glycosylases. The monofunctional glycosylases utilize an activated water molecule to initiate nucleophilic attack the C1’ carbon of the N-glycosidic bond, liberating the damaged base in a single-step process (Figure 9). In contrast, the bifunctional glycosylases instead use an amine group (such as the epsilon amino group found on lysine) to attack C1’; this generates a covalent enzyme-substrate complex. This Schiff base intermediate is resolved by cleavage of the DNA backbone on the 3’ side (β-elimination). Some glycosylases, such as the NEILs, additionally cleave on the 5’ side as well (β, δ-elimination) (reviewed in (Fromme et al., 2004) (Kim, 2012 #498)).

DNA glycosylases are typically separated into four phylogenetic “superfamilies”, named for their bacterial (or eukaryotic) representatives: 1) the uracil DNA glycosylases, 2) the alkyladenine DNA glycosylase (AAG) superfamily, 3) the Nth superfamily, and 4) the Fpg/Nei family. In general, the bifunctional glycosylases of the Nth and Fpg/Nei superfamilies are larger and more complex than the small, single domain, monofunctional glycosylases of the UDG & AAG superfamilies. The focus in this section will be primarily on those DNA glycosylases (bacterial and human) that recognize oxidative
damages, therefore I will largely be highlighting members of the second two superfamilies (for a list of the major known human glycosylases, refer to Table 2).

The largest of the four superfamilies is the Nth, named for *E. coli* endonuclease III (eNdonuclease THree → “Nth”) (Thayer et al., 1995). Members of this family share a helix-hairpin-helix (HhH) motif. Bifunctional glycosylases typically contain a catalytic Lys residue within this motif that is necessary for AP lyase activity, although the HhH motif itself is implicated in DNA binding (Nash et al., 1996) (Doherty et al., 1996). Nth superfamily glycosylases also contain a catalytic Asp residue and a glycine/proline–aspartate (GPD) motif that is also involved in catalysis. Because of these structural features, the Nth superfamily is sometimes referred to as the “HhH-GPD” superfamily. The major human glycosylases within this family that recognize oxidized bases are endonuclease three-like 1 (NTH1 or NTHL1), 8-oxoguanine glycosylase 1 (OGG1), and MutY homolog (MUTYH) (Eisen & Hanawalt, 1999).

Members of the Fpg/Nei family (named for formamidopyrimidine and endonuclease VIII DNA glycosylase), also shares conserved domains/motifs, including either a proline–glutamate helix or a helix–two turn–helix motif, and a zinc finger (Zharkov et al., 2003). Fpg/Nei glycosylases also share an N-terminal proline that acts as the catalytic nucleophile. There are no known Fpg (also known as MutM) homologs in humans, but there are Nei-like (NEIL) glycosylases. To date, three have been discovered (Hazra, Izumi, et al., 2002) (Bandaru et al., 2002) (Hazra, Kow, et al., 2002). However,
the human NEIL glycosylases lack some of these structural traits. NEIL1 contains a structurally similar “zincless finger” motif instead of zinc finger (Doublie et al., 2004) and NEIL3 contains an N-terminal valine as a catalytic nucleophile instead of proline (Liu, Bandaru, et al., 2010).

Glycosylases in eukaryotes often have features not found in prokaryotes, which are thought to allow them to act in a coordinated manner with other repair proteins. For example, eukaryotic glycosylases tend to bind AP sites tightly, leading to poor turnover and product inhibition (for reviews, see (Wilson & Kunkel, 2000) (Prasad et al., 2011)). This effectively sequesters the AP site, which itself is a replication stalling damage, until the next enzyme comes in to process the intermediate, in a process known as substrate handoff. APE1, discussed later, can displace glycosylases from the AP site intermediate, serving the dual purpose of: 1) enhancing the activity (through increased turnover) of the glycosylase and 2) initiating the next step of repair. A second notable difference, also attributable to substrate handoff, is the existence of “disordered” regions within the mammalian DNA glycosylases as well as other DNA repair proteins. These regions tends to carry positive or negatively charged amino acids (Arg, Lys, Glu, Asp) and are implicated in numerous protein-protein interactions (reviewed in (Hegde et al., 2012)).

**Human endonuclease three-like 1 (NTHL1 or NTH1)** is an Nth homolog that recognizes thymine glycol (Tg), as well as DHU, 5-OHC, 5-OHU, FapyG and hydantoins (Hilbert et al., 1997) (Aspinwall et al., 1997) (Sarker et al., 1998). NTHL1 also
recognizes Tg:G mispairs arising from oxidation of 5-methylcytosine, although this particular mispaired lesion can additionally be recognized by human monofunctional glycosylases thymine DNA glycosylase (TDG) and methyl-CpG-binding domain protein 4 (MBD4) (Yoon et al., 2003). NTHL1 is generally thought to be present in all cell types and throughout the cell cycle, although one study in HeLa cells suggests that it may also be upregulated during S-phase (Luna et al., 2000). hNTHL1 contains a disordered N-terminal tail that facilitates hNTHL1 dimerization and may be implicated in protein-protein interactions. Both removal of the tail and dimerization increase enzymatic turnover (Liu & Roy, 2002) (Liu et al., 2003).

Early NTHL1 mouse knockouts displayed no obvious phenotypic defects, even when mice were allowed to age for two years (Takao et al., 2002) (Ocampo et al., 2002). Ocampo et al. extracted tissue from the Nth1 null mice and discovered that these extracts could still process Tg:G and Tg:A lesions, demonstrating that the loss of a single glycosylase can often be compensated for by the overlapping substrate specificity of other glycosylases. However, NTHL1 appears to be the only glycosylase able to remove Tg lesions from mouse mitochondria, as knockout ablates this particular repair activity (Karahalil et al., 2003). Examination of telomeric repeats in Nth1 null mouse tissue revealed telomeric defects, suggesting a role for NTHL1 in telomere maintenance. These cells not only contained higher levels of lesions recognized by NTHL1, but these lesions were also subject to slower repair. As well, bone marrow cells exhibited telomere shortening, increased recombination, and DNA damage foci formation when exposed to
increased oxygen levels. When subjected to replicative stress, hematopoietic tissue demonstrated telomere loss and apoptosis (Vallabhaneni et al., 2013).

NTHL1 contains an iron-sulfur (Fe-S) cluster motif whose function still remains somewhat mysterious (Kuo et al., 1992) (Thayer et al., 1995). The cluster is located far from the active site in the crystal structure of bacterial Nth, arguing against a role in catalysis. Bacterial MutY glycosylase (in humans, MUTYH) also contains a Fe-S cluster that is essential to DNA binding and activity (Porello et al., 1998). In both Nth and MutY, the redox potential of the Fe-S cluster shifts upon DNA binding (Boal et al., 2005). These observations, and the fact that the π-orbitals of the DNA base pair stack can transmit a charge, led the Barton group to propose a novel “search” function for the Fe-S cluster. Briefly, two reduced glycosylases (in this case NTHL1) bind a length of DNA and transmit an electron through it. If there is no perturbing base stacking chemistry (which can be as little as a single base pair mismatch), the electrons travel the length of the DNA and re-reduce the glycosylases, allowing the enzymes to dissociate from the DNA and diffuse to another region. However, any DNA lesion that disrupts the base stack will block charge transport. In this scenario, both glycosylases remain oxidized and bound to DNA, and slide along the helix until the damage is found (Grodick et al., 2015). While this remains theoretical and undemonstrated in vivo, it is a very interesting model. Interestingly, this charge transfer phenomenon is able to occur through a nucleosomal substrate with similar efficiency as naked DNA (Nunez et al., 2002).
8-Oxoguanine glycosylase 1 (OGG1). While many glycosylases are strongly conserved between prokaryotes and eukaryotes, there are exceptions. Formamidopyrimidine DNA glycosylase (Fpg or MutM) in *E. coli* has no mammalian homolog. Instead, the “functional” homolog in mammals that recognizes 8-oxoG (as well as FapyG) paired with C is OGG1, which belongs to the Nth superfamily (Lu et al., 1997) (Hazra et al., 1998) (Krokan et al., 2000). If an 8-oxoG lesion persists through replication and mispairs with adenine, the glycosylase MUTYH can remove nascent, mispaired adenine, before or during replication (Hazra et al., 2001). This second chance for repair following replication makes 8-oxoG unique among the oxidative damages.

Homozygous *Ogg1*−/− mice are viable, but contain increased 8-oxoG levels as well as ~10-fold elevated G:C to T:A spontaneous transversions (Klungland et al., 1999). *MutY*−/−*Ogg1*−/− double knockout mice exhibit increased tumor formation, particularly lung and lymphoma, whereas single knockouts of either OGG1 or MUTYH do not (Xie et al., 2004). As well, OGG1 mutations and polymorphisms are found in several types of cancer tumors. (Chevillard et al., 1998) (Blons et al., 1999) (Dherin et al., 1999).

OGG1 has poor AP lyase activity and is product inhibited by an AP site. Consequently, it is readily stimulated by APE1 (discussed in the next section). In one form of “stimulation”, APE1 enhances strand cleavage on DNA by bypassing OGG1’s lyase activity altogether with its own. However, APE1 also directly stimulates OGG1 glycosylase activity by displacing OGG1 from AP sites. This has been demonstrated by
the stimulation of OGG1 by catalytically dead APE1 mutants, or by WT APE1 in reactions lacking Mg\(^{2+}\) (necessary for APE1 catalytic activity) (Hill et al., 2001). In the same study, APE1 does not appear to physically interact with OGG1 via co-immunoprecipitation, further indicating that activity stimulation is through displacement. OGG1 can also be stimulated via AP site turnover by fellow glycosylases. For example, although NEIL1 and OGG1 do not physically interact, NEIL1 can stimulate OGG1 activity and bypasses its lyase activity (Mokkapati et al., 2004). However, OGG1 cannot stimulate NEIL1 in return; this is evidenced by NEIL1’s higher affinity for non-hydrolysable AP-analog site (tetrahydrofuran).

The Seeberg group demonstrated that nuclear OGG1 localization is both cell-cycle and phosphorylation dependent (Dantzer et al., 2002). During S-phase, OGG1 is localized to the nucleoli, possibly as a sequestration mechanism while MUTYH acts as the replication-associated glycosylase to remove 8-oxoG:A. During interphase, however, OGG1 is localized to chromatin and the nuclear matrix, and during mitosis, it localizes to condensed chromatin. This localization appears to depend on phosphorylation at Ser326; consequently, the Ser326Cys polymorphism (associated with several tumor and cancer types) is mislocalized at each cell cycle stage. Co-immunoprecipitation with OGG1 suggests that this phosphorylation is carried out by protein kinase C (PKC), although this has not yet been demonstrated definitively in vivo (Luna et al., 2005).
**Nei endonuclease VIII like 1-3 (NEIL1-3) DNA glycosylases.** The human Nei homologs, NEIL1-3, are bifunctional glycosylases with broad substrate specificities. NEIL1 removes pyrimidine lesions such as Tg, 5-OHU, and 5-OHC, but also recognizes purine lesions such as FapyA and FapyG (Bandaru et al., 2002) (Krishnamurthy et al., 2008). NEIL2 and 3 have similar substrate specificities to NEIL1. Unique among the glycosylases, the NEILs can recognize base damages located within single-stranded DNA sequences, including bubble substrates, fork substrates, and quadruplex DNA. While NTHL1 and OGG1 seem to function primarily as the “housekeeping” glycosylases for oxidative lesions in humans, the NEILs likely have more specialized roles.

NEIL1 is a likely candidate for the major replicative glycosylase associated with repair of oxidized bases. Its sequence contains a 100 residue, C-terminal disordered tail which interacts with several replication proteins, including FEN-1, PCNA, RPA, and pol δ (Das et al., 2007) (Dou et al., 2008) (Hegde, Theriot, et al., 2008) (Theriot et al., 2010) (Hegde et al., 2013). Additionally, the expression of NEIL1 is cell cycle dependent, increasing with S-phase. The Mitra group has proposed a mechanism wherein NEIL1 binds to an oxidized base in ssDNA, but is prevented from acting by RPA (thus avoiding immediate strand cleavage). This binding blocks pol δ, triggering fork regression, which enables the DNA containing the damaged base to reanneal to its complement, in a reaction promoted by the Werner helicase (WRN). NEIL1 can then excise the base. Repair is likely completed by long-patch BER due to proximity to the replication machinery (Dutta et al., 2015). Both nullizygous and homozygous NEIL1 mice are viable
with a normal phenotype at birth. However, interestingly, some males develop metabolic syndrome over time. This is possibly due to elevated rates of DNA damage and deletion mutations within null mice mitochondria (Vartanian et al., 2006).

Overlapping substrate specificities suggest that NEIL2 may be able to substitute for NEIL1 function in replication. However, NEIL2’s main role seems to be in transcription-associated BER. This inference is based on interaction between NEIL2 and RNA polymerase II and heterogeneous nuclear ribonucleoprotein U (hnRNP-U) (Banerjee et al., 2011). Neil2−/− mice display no obvious phenotype, but when the mice are allowed to age, they accumulate oxidative damages, particularly within actively transcribed regions of the genome, further supporting a role for NEIL2 in transcription-coupled BER (Chakraborty et al., 2015).

The expression of NEIL3 is dependent upon cell type, and in humans appears to be restricted to highly proliferating cells as well as certain types of cancer cells (Hildrestrand et al., 2009) (Regnell et al., 2012) (Kauffmann et al., 2008). Knockout mouse models have revealed a possible link between NEIL3 and neurogenesis. While Neil3−/− mice are viable, they exhibit learning and memory defects as well as high levels of neuropathology following hypoxia–ischemia (Torisu et al., 2005) (Sejersted et al., 2011) (Regnell et al., 2012). There is also evidence for NEIL3 as a glycosylase that recognizes telomeric oxidative damage in vivo. Recent studies have indicated that NEIL1 but particularly NEIL3 can recognize base lesions within a telomeric sequence context.
(TTAGGG), with NEIL3 having a strong preference for the sequence. G-rich telomeric sequences (as well as some promoter sequences) can form a secondary, four-stranded structure known as quadruplex DNA [Williamson, 1989 #499] (Smith & Feigon, 1992). Both NEIL1 and NEIL3 can process lesions from certain positions within this structure. While no known glycosylase can excise 8-oxoG from quadruplex DNA, both NEIL1 and 3 can remove guanidinohydantoin (Gh), and spiroiminodihydantoin (Sp), which are further oxidation products of 8-oxoG. In addition, they can also recognize some hydantoin. However, only NEIL3 can remove the major pyrimidine damage Tg from quadruplex DNA (Zhou, Liu, et al., 2013) (Zhou, Fleming, et al., 2015).

1.3.6. AP Endonuclease 1 (APE1)

*E. coli* possess two types of apurinic/apyrimidinic (AP) endonucleases: exonuclease III (Xth) and endonuclease IV (Nfo). Inactivation of these genes results in increased sensitivity to certain DNA damaging agents, but did not impair viability (Cunningham et al., 1986). Homologs of both of these endonucleases can be found in lower eukaryotes such as yeast, and appear to be similarly dispensable to survival. In contrast, mammals have only one AP endonuclease, APE1 which is in the Xth family. (A putative Xth-type APE2 was cloned and characterized, but its function appears unrelated to BER and it cannot complement APE1 mutants (Tsuchimoto et al., 2001)). APE1 is essential for both short and long patch repair, and accounts for ~95% of endonuclease activity in the cell (Demple & Harrison, 1994) (Doetsch & Cunningham, 1990). Multiple studies have corroborated the embryonic lethality of mouse APE1 knockout
Heterozygous mutants are viable and develop normally, however, primary cell isolates from these mice exhibit increased sensitivity to oxidizing agents (Meira et al., 2001). Conversely, APE1 overexpression occurs in many cancer tumors, and is correlated with drug resistance and poor prognosis (reviewed in (Mitra et al., 2007)).

The major role of APE1 in BER is its endonuclease activity, which cleaves the DNA backbone 5’ to the abasic site, resulting in a clean 3’ hydroxy (OH) group and a 5’ deoxyribophosphate (dRP) moiety. This dRP group is then removed by pol β using a separate dRP lyase activity before the resulting gap is filled. However, the intrinsic AP lyase activity of bifunctional glycosylases will result in some kind of 3’ blocking group: either a α,β-unsaturated aldehyde (3’PUA) or a phosphate group. Either group must be removed to complete subsequent repair steps, as polymerase activity requires a 3’OH, while ligation requires the 3’OH as well as a 5’P. The 3’-PUA (created by hNTHL1 and hOGG1) is removed by APE1 through its endonuclease activity. APE1 also has the ability to excise blocking groups such as 3’-phosphates, 3’-phosphoglycolate esters, and 3’-deoxyribose fragments. However, APE1 does not appear to have any activity on the 3’P generated by AP lyase activity of human NEIL1-3 (or sometimes by ROS). Instead, polynucleotide kinase (PNK) removes the 3’P; although there also exist other human enzymes that can remove a 3’P (refer to (Mitra et al., 2007) and (Li & Wilson, 2014) reviews). Surprisingly, APE1 can also recognize several base damages without the need of glycosylase recognition, including: 5,6-dihydro-2’-deoxyuridine, 5,6-
dihydrothymidine, 5-hydroxy-2'-deoxyuridine, alpha-2'-deoxyadenosine and alpha-thymidine adducts. This activity has been dubbed the “nucleotide incision repair” (NIR) subpathway of BER (Gros et al., 2004).

The crystal structure of APE1 reveals a few commonalities with the glycosylases (Mol et al., 2000) (Freudenthal et al., 2015) (Figure 10). APE1 binds the minor groove in a similar manner, flipping the AP site out of the DNA helix while bending the DNA ~35°; the DNA is stabilized by four loops and one α-helix. Endonuclease IV, which is structurally unrelated to APE1, does this as well, indicating that this may be a conserved property of AP endonucleases (Hosfield et al., 1999). Strand cleavage at the AP site is effected through hydrolysis, but the exact mechanism of APE1 is still under controversy. Key active site residues include Glu96, which coordinates a divalent metal ion (Mg$^{+2}$), and Asp210/His309, which are necessary for hydrolysis. While APE1 requires Mg$^{+2}$ to function, it is not required to bind DNA (reviewed in (Li & Wilson, 2014)).

The lyase activities of the DNA glycosylases vary, but are generally worse than APE1. APE1 endonuclease activity is significantly more robust, to the point where many bifunctional glycosylases (particularly OGG1) act essentially as monofunctional glycosylases in vivo (Mitra et al., 2007). As well, APE1 has a tighter binding affinity for AP sites, allowing it to displace product-bound glycosylases. In a manner similar to the described above for OGG1, APE1 stimulates turnover of multiple other glycosylases, including MUTYH, NTHL1, UNG2, and TDG (reviewed in (Jacobs & Schar, 2012)).
APE1 contains a redox regulatory region within its N-terminal domain that is not found in *E. coli* Xth. This region can be mutated in mouse embryonic fibroblasts without affecting cell viability (only the AP endonuclease function is essential), but it allows APE1 to act as a redox signaling protein, making APE1 the only known DNA repair factor to control expression of other repair proteins in this manner. APE1 was actually independently identified and cloned as redox factor 1 (Ref-1) at about the same time that it was discovered as APE1. In 1992, it was discovered to stimulate (Fos/Jun) transcription factor complex DNA-binding activity to the gene of activator protein-1 (AP-1) via redox (Xanthoudakis & Curran, 1992). APE1 has also been demonstrated to have similar activity on additional transcription factors NF-κB, Myb, p53, and members of the ATF/CREB family (reviewed in (Kelley et al., 2012)).

The exact redox mechanism of APE1 is not well understood. APE1 reduces a conserved cysteine residue within the DNA binding domain of the target protein, but lacks the C-X-X-C motif common to thioredoxin and other redox proteins (Hirota et al., 1997). Therefore, it is uncertain how APE1 creates the disulfide bond intermediate necessary for redox reaction. The crystal structure of APE1 does not indicate two cysteines within correct proximity of each other for bond formation (Gorman et al., 1997) (Beernink et al., 2001) (Georgiadis et al., 2008). It has been hypothesized that APE1 must undergo a conformational change in order to expose a buried cysteine to form the necessary disulfide (Luo et al., 2012).
1.3.7. DNA Polymerase β (pol β)

A member of the X family of polymerases, pol β is the smallest known eukaryotic polymerase. It is a single subunit, containing an 8 kDa N-terminal domain with lyase activity and a C-terminal 31 kDa polymerase domain. The lyase domain allows pol β to remove the 5’-dRP group left behind by APE1 following strand incision. Pol β is a relatively low fidelity polymerase, due to the absence of a proofreading domain. While mouse knockout of pol β is embryonic lethal, cell lines are viable but sensitive to alkylating agents (Gu et al., 1994) (Sobol et al., 1996). Pol β is implicated as the polymerase involved in short-patch BER. It can function in long-patch repair (discussed later), where strand displacement is required, although more commonly synthesis following the initial base incorporation is probably carried out by the replicative polymerases pol δ/ε (reviewed in (Yamtich & Sweasy, 2010)).

Polymerases domains are commonly described as a “right hand” structure with a thumb, palm, and fingers. The first step of nucleotide incorporation begins with pol β binding a DNA substrate, preferably a single nucleotide gap with 5’P and 3’OH ends, and bending the DNA ~90°. At this point, the polymerase is in an “open” configuration. The fingers then subsequently close into an active configuration upon preferential binding of the correct dNTP. This conformational change aligns the 3’ oxygen on the primer end to attack the α-phosphate of the incoming dNTP. In human pol β, reaction requires catalytic residues Asp190, Asp192, Asp256 as well as two coordinated metal ions. Following
nucleotide incorporation, pol β likely “opens” again, releasing the DNA substrate and pyrophosphate (see (Beard & Wilson, 2014) for a detailed review).

While pol β appears to have many interaction partners, its interaction with XRCC1 is probably the best-studied. The palm and thumb domains of pol β bind the N-terminal domain of XRCC1 (Dianova et al., 2004) (see Figure 12). As XRCC1 can also bind gapped DNA, it may serve to recruit to pol β to a gap or SSB and thus serve to protect the repair intermediate in the interim. Conversely, pol β may provide a similar function, recruiting XRCC1 (and correspondingly, LigIIIα – see the next section for detail) to the nick product following nucleotide incorporation (reviewed in (Yamtich & Sweasy, 2010))

1.3.8. Ligase IIIα & XRCC1 (LigIIIα-XRCC1)

With the exception of some viral ligases, all known DNA ligases completely encircle their DNA target. All ligases contain an oligonucleotide/oligosaccharide binding (OB)-fold domain, which binds DNA via the minor groove, as well as a nucleotidyltransferase (NTase) domain, joined by a flexible linker. Additionally, the three mammalian DNA ligases (Ligase I, III and IV) contain an N-terminal DNA-binding domain (DBD) that can bind the minor groove of DNA and enhance nick joining activity.

The final step of BER is ligation of a nick containing 3’OH and 5’P ends by LigIIIα. Eukaryotic DNA ligases all use ATP (whereas some bacterial, archaeal, and viral
Ligases use NAD\(^+\)). In the first step of ligation, ATP is hydrolyzed, leaving AMP covalently bound to an active site lysine residue. In the second step, AMP is transferred from lysine to the 5’P DNA end. This facilitates the third step, which entails a nucleophilic attack by the 3’OH on the 5’-AMP, resulting in phosphodiester bond formation and release of AMP (see Figure 11) (reviewed in (Ellenberger & Tomkinson, 2008)).

LigIII\(\alpha\) is one of four splice variants in humans. The first two are nuclear and mitochondrial forms of LigIII\(\alpha\), while the second two are nuclear and mitochondrial forms of LigIII\(\beta\). LigIII\(\beta\) appears to only be expressed in germ cells. However, the nuclear form of LigIII\(\beta\) is the one that is most often studied, due to the fact that LigIII\(\alpha\) is unstable without XRCC1 (discussed below).

LigIII is unique among the mammalian ligases for its additional, N-terminal zinc finger (ZnF). The ZnF enhances DNA binding and nick sensing, and is indispensable to LigIII’s blunt-end ligation activity. The ZnF possibly binds DNA first, but is then displaced by the DBD and OB-fold domains upon conformational change into an active state (Cotner-Gohara et al., 2010) (Kukshal et al., 2015). Finally LigIII\(\alpha\) (but not LigIII\(\beta\)) contains a BRCA1 Carboxyl Terminal (BRCT) domain which mediates its interaction with XRCC1 (Taylor et al., 1998).
X-ray repair cross-complementing protein 1 (XRCC1). As its name implies, XRCC1 was originally discovered for its ability to complement the phenotype of mutant CHO cell lines (EM7, EM9, EM-C11, and EM-C12), which display enhanced sensitivity to multiple DNA damaging agents, and are deficient in single-strand break repair (SSBR) (reviewed in (Caldecott, 2003)). When His-tagged XRCC1 was first expressed and purified from EM9 cells, LigIIIα was found to co-purify (Caldecott et al., 1994). This observation suggested that the two proteins strongly interact in vivo. Furthermore, loss of XRCC1 appears to be correlated with a significant drop in LigIIIα activity, suggesting that XRCC1 is required for LigIIIα stability. It has been hypothesized that this dependence of LigIIIα on XRCC1 is due to the lack of a nuclear localization signal (NLS). Thus, without XRCC1 for nuclear targeting, LigIIIα becomes subject to proteolytic degradation in the cytosol. Loss of XRCC1 in mice is embryonic lethal, although this could be due to combined loss of both XRCC1 and nuclear LigIIIα activity, as LigIIIα<sup>−/−</sup> mice do not survive either (Tebbs et al., 1999) (Puebla-Osorio et al., 2006). However, more recent studies suggest that the essential function of LigIIIα is related to its mitochondrial splice variant (which does not require XRCC1 for stability), and not its role in BER/SSBR, for which LigI can substitute (Shokolenko et al., 2013).

XRCC1 is a 633 residue, 85 kDa scaffolding protein with no known enzymatic activity. XRCC1 interacts with a number of proteins involved in BER, NER, replication, and other pathways (reviewed in (London, 2015)). XRCC1 also has significant DNA binding affinity for nicks and 1 nt gaps (K<sub>D</sub>’s ~65 nM and ~34 nM, respectively) (Mani et
XRCC1 contains an N-terminal domain and two BRCT domains. The N-terminal domain is responsible for binding DNA and pol β, while the BRCTI and BRCTII domains interact with PARP-1 and LigIIIα, respectively. Additionally, the protein contains two intrinsically disordered regions located between NTD-BRCTI and BRCTI-BRCTII, which are implicated in binding/interacting with multiple proteins, including multiple glycosylases, APE1, PNPK, and PCNA (see Figure 12).

This scaffolding activity of XRCC1 may allow it to coordinate DNA repair based upon the nature and context of the damage. For example, during replication, XRCC1 co-localizes with PCNA and UNG2 foci (UNG2 is a replication associated glycosylase). Endogenous, non-S-phase XRCC1 foci that lack PCNA and UDG2 also exist. Immunoprecipitation and gel filtration of HeLa cell extracts stably expressing XRCC1-EYFP revealed complexes of various sizes with different DNA repair activities, implicating XRCC1 as a player in multiple DNA repair complexes. (Hanssen-Bauer et al., 2011).

Several post translational modifications of XRCC1 have been found to influence DNA repair. Checkpoint kinase 2 (Chk2), which is activated by ataxia telangiectasia-mutated (ATM) upon DNA damage, phosphorylates XRCC1 on Thr288. In addition to phosphorylating XRCC1, Chk2 and XRCC1 interact directly; this interaction increases with the addition of DNA damaging agents. A XRCC1 point mutation that mimics phosphorylation of Thr288 demonstrated that this phosphorylation increases the
affinity of XRCC1 for glycosylases MPG and UNG2, but not for downstream BER proteins. Phosphorylation also seem to play a role in methyl methanesulfonate-induced recruitment of XRCC1 to chromatin. Blocking phosphorylation by either XRCC1 mutation of Chk2 or siRNA of Chk2 reduced efficiency of BER (as measured by PARP-1 activity, SSB formation, and cell survival) (Chou et al., 2008).

The essential protein kinase casein kinase 2 (CK2) phosphorylates XRCC1 at multiple sites, including Ser518, Thr519, and Thr523. This phosphorylation has been demonstrated to block ubiquitination and subsequent degradation of XRCC1 in the cytoplasm before it can be transported into the nucleus (Parsons et al., 2010). As well, phosphorylation by CK2 allows for XRCC1 to interaction with the forkhead associated (FHA) domain of both APTX and PNK (Clements et al., 2004).

1.3.9. BER Accessory & Interacting Proteins

Poly ADP-ribose polymerase 1 (PARP-1) influences multiple cellular pathways, making its function with respect to DNA repair ambiguous. PARP-1 is also associated with chromatin architecture, gene expression, and apoptosis, making it difficult to determine which DNA repair effects are direct, and which may arise indirectly through PARP-1’s influence on associated pathways. PARP-1 is a 113 kDa nuclear protein that contains an N-terminal DNA Binding Domain (DBD) as well as three zinc finger (ZnF) motifs, two of which are implicated in DNA binding. These regions allow PARP-1 to
function as the initiating single-strand break sensor in the SSBR subpathway of BER. Nicked DNA visualized by atomic-force microscopy appears intrinsically kinked at a ~120° angle; binding of PARP-1 to the nick bends the DNA back to a ~100° angle. As well, PARP-1 can bind DNA ends (Le Cam et al., 1994).

PARP-1 stimulates SSBR in cell free extracts, but it is not required for efficient repair. Instead, it appears that PARP-1 is required primarily for recruitment of other repair factors in vivo. PARP-1’s “PARylation” activity creates negatively-charged poly ADP-ribose (PAR) chains and branches using NAD+, and then transfers them to histones and other proteins, including itself. Auto-PARylated PARP-1 binds to both XRCC1 and LigIIIα, and is necessary for XRCC1 foci formation in MEFs (reviewed in (Caldecott, 2003)).

The role of PARP-1 in BER is more subtle, as the single strand break induced by glycosylase or APE1 can likely be handed off directly to the next player, bypassing the need for PARP-1. However, PARP-1 deficient MEFs are sensitive not only to strand break-inducing agents such as IR, but also to alkylating agents (Trucco et al., 1998). Another study found incomplete repair of an 8-oxoG substrate in double knockout Parp-1−/−/Polβ−/− 3T3 cells, whereas corresponding single knockouts were repair-competent (Le Page et al., 2003). Conversely, a conflicting study using a different methodology and a different MEF cell line found no sensitivity to alkylating agents (Vodenicharov et al., 2000). One proposed mechanism of PARP-1-enhanced BER is through stimulation of
long patch repair. Evidence for this includes the reduced expression of pol β and FEN1 in PARP-1 deficient cells, as well as strand displacement stimulation in the presence of pol β and FEN1 (Sanderson & Lindahl, 2002) (Prasad et al., 2001).

**Tyrosyl-DNA phosphodiesterase 1 (TDP1).** Human topoisomerase I (TOP1), which relaxes over-twisted DNA, generates a SSB as a reaction intermediate, with Tyr723 covalently binding to the 3’ end of DNA (Pourquier & Pommier, 2001). TDP1 was named for its ability to cleave this covalent bond in the event that TOP1 fails to release and re-ligate the break (Interthal et al., 2001). Following tyrosine removal, the resulting 3’P left by TDP1 can then be removed by PNK, allowing for ligation by LigIIIα. Mutation of TDP1 in humans is associated with the rare spinocerebellar ataxia with axonal neuropathy 1, while knockout of TDP1 renders vertebrate cells hypersensitive to anti-cancer drugs, ROS, and alkylating agents (Murai et al., 2012). TDP1 can also remove glycolate from 3’-phosphoglycolate overhangs, a common product of oxidized strand breaks or the anti-cancer drug bleomycin (Inamdar et al., 2002). There are a number of additional covalently bound 3’ substrates recognized by TDP1, including protein-DNA adducts, oxidized nucleotides, and chain terminators. These products are sometimes the result of chemotherapeutics, making TDP1 an attractive target for development of small molecule inhibitors (reviewed in (Comeaux & van Waardenburg, 2014)).
TDP1 interacts with both XRCC1 and LigIIIα, suggesting a role for it in SSBR and/or BER. Besides being available for processing of certain unligatable 3’ ends, TDP1 is also able to remove 3’-dRP lesions and cleave AP sites. In yeast, TDP1 can substitute for the function of APE1, although it prefers either a bubble-substrate AP site, or an AP site located across from a bulky adduct (Nilsen et al., 2012) (Lebedeva et al., 2012). Double TDP1/APE1 knockdowns rendered human MRC5 cells acutely hypersensitive to alkylating agents, more so than single knockdown of either enzyme. Thus, TDP1 may be able to substitute for APE1 in mammalian cells as well (Alagoz et al., 2014).

**Aprataxin (APTX)** is a member of the histidine triad (HIT) superfamily of nucleotide hydrolases/transferases. Mutation of APTX is associated with the common recessive spinocerebellar ataxia: ataxia oculomotor apraxia 1 (AOA1). The likely physiological target is adenosine monophosphate (AMP) covalently bound to 5’ DNA ends. This 5’-AMP group is an intermediate generated during abortive ligation, and can occur when ligase attempts to ligate non-ligatable ends, such as those at “dirty breaks”. By removing the blocking AMP moiety, ligase can re-attempt the ligation event after the ends have been processed correctly. APTX can also remove phosphate and phosphoglycolate moieties from the 3’ end, which can arise as a result of IR and ROS (Takahashi et al., 2007).

APTX deficiencies cause defects in short patch BER, but the mechanism behind this is surprisingly not due to accumulation of 5’-AMP intermediates. Instead, it is due to
the depletion of non-adenylated LigIIIα. Ligase bound to AMP (LigIIIα-AMP) cannot catalyze the final nick-sealing step. Therefore, when LigIIIα aborts and dissociates from 5′-AMP DNA, it will hydrolyze readily available ATP and reform LigIIIα-AMP, and is now unable to return to seal the nick. (Reynolds et al., 2009). Additionally, a 5′-adenylated-deoxyribose phosphate (5′-AMP-dRP) can form if ligase attempts to ligate a strand before the dRP group is removed by pol β. APTX can remove the 5′-AMP, leaving the 5′dRP to allow for another attempt at dRP cleavage (Caglayan et al., 2015). In both cases, LP-BER (specifically FEN1 activity) can compensate for APTX deficiency in cell extracts.

1.3.10. Choice Between BER Subpathways

The short-patch BER (SP-BER) enzymes APE1 (or PNK), pol β, and LigIIIα-XRCC1 are constitutively expressed, making it similarly efficient in both proliferating and non-proliferating cells. Due to the need for replicative enzymes, long-patch BER (LP-BER) appears to be the minor pathway in terminally differentiated human cells, with SP-BER predominating (reviewed in (Krokan & Bjoras, 2013)). Long-patch BER (LP-BER) requires APE1/PNK as well, but subsequent enzymes are associated with DNA replication, including pol δ/ ε, FEN1, LigI PCNA, and RFC. Pol β can participate in LP-BER as well, and is the preferred polymerase in non-proliferating cells (Akbari et al., 2009). Repair by replication-associated glycosylases such as UDG and NEIL1 may preferentially occur via LP-BER due to the proximity of the replicative machinery.
The DNA glycosylases that initiate BER may also play a role in determining which BER subpathway is taken. Fortini et al. incubated HeLa cell extracts with circular plasmids containing base lesions targeted by either 3-methyladenine DNA glycosylase (ANPG protein) or OGG1. By limiting reactions to only one type of radiolabeled dNTP (corresponding to the 1st or 2nd dNTP incorporated), the authors were able to differentiate between SP-BER (in which only the first nt is incorporated) and LP-BER (where the 2nd nt will be incorporated). ANPG substrates were subject to both LP-BER and SP-BER, whereas OGG1 substrates exclusively underwent SP-BER (Fortini et al., 1999).

Switching from SP-BER to LP-BER may occur when pol β encounters a non-dRP 5’ blocking group that cannot be processed. For example, 2-deoxyribonolactone (dL) is a C1’-oxidized AP site that can arise from damaging agents such as IR or UV. This reactive moiety can generate protein-DNA cross-links with various BER enzymes. Cross-links created with glycosylases and AP endonucleases appear to resolve themselves efficiently, whereas pol β cross-linking persists, stalling SP-BER. However, LP-BER can bypass this block due to FEN1 excision of 5’-dL on a displaced oligonucleotide flap (Sung et al., 2005).

High mobility group box 1 (HMGB1) is an abundant architectural protein with a variety of cellular functions, and potentially plays a role in facilitating pol β-dependent LP-BER. HMGB1 binds damaged DNA, particularly DNA crosslinks, as well as
noncanonical DNA structures (e.g. supercoiled DNA, Z-DNA, and single-stranded DNA) (Bidney & Reeck, 1978) (Hamada & Bustin, 1985). Binding by HMGB1 induces further DNA distortion and can disrupt nucleosome structure by bending linker DNA. HMGB1 can compete for chromatin binding with H1, generating a less compact chromatin fiber and increasing accessibility to chromatin regulatory factors (Catez et al., 2004). The protein has been implicated in NER, MMR, DSBR, and BER, although the mechanisms surrounding these interactions are not well understood (reviewed in (Lange & Vasquez, 2009)). A role for HMGB1 in BER was found during an initial screen for proteins with dRP lyase activity. While its dRP lyase activity is far weaker than pol β, subsequent pull-downs identified interactions with APE-1, FEN1, and pol β. Under in vitro conditions, HMGB1 enhanced activity of APE1 and FEN1, while mouse cell HMGB1 knockout was resistant to the alkylating agent methyl methanesulfonate. This demonstrated that HMGB1 stimulates the production of BER strand-break intermediates (Prasad, Liu, et al., 2007). A later study proposed that the weak dRP lyase activity of HMGB1, combined with its cellular abundance, causes it to act as binding competition for pol β and therefore inhibit SP-BER, favoring LP-BER instead through its stimulation of FEN1 flap endonuclease activity (Liu, Prasad, et al., 2010).
1.4. DNA Repair in the Context of Chromatin

The hydroxyl radical reacts with DNA to generate oxidative base damages. The minor grooves within nucleosomes, however, are alternately protected or exposed in five base pair intervals, providing a ~2 to 3-fold level of protection from ROS (Hayes et al., 1990) (Luger et al., 1997). As well, linker histone and non-histone, chromatin associated proteins offer additional an additional layer of protection from oxidative damage (reviewed in (Cannan & Pederson, 2016)). This would imply that oxidative DNA damage would be located primarily within non-nucleosomal DNA such as linker DNA. However, (Enright et al., 1996) demonstrated that this is not the case. Thus, nucleosomes may provide some protection from ROS, but not total protection. A reasonable hypothesis follows: if damage can occur within chromatin, then it could be repaired within chromatin as well. Among our most fundamental questions are whether or not the BER enzymes, by themselves or in conjunction with other proteins, can recognize and remove damages from within a nucleosome. Our lab and others have studied BER using an in vitro system that utilizes nucleosomes containing specifically positioned DNA damages. The following sections highlight some of the insights learned from this system.

1.4.1. Systems Used to Study BER in Nucleosomes

An in vitro system for investigating DNA repair on the nucleosome typically involves isolation of a homogenous histone octamer as well as a DNA sequence with
well-defined translational and rotational positions. It is important to understand the commonly used methods for preparing such reagents, as they can be a source of variation between different labs and studies. For a recent review of the following methods, refer to (Taylor, 2015).

Transcriptionally dormant chicken erythrocytes can provide a substantial amount of fairly homogenous histone octamer. Single nucleosomes can be isolated via micrococcal nuclease digestion followed by size-exclusion chromatography; alternatively, the histones alone can be extracted and subsequently reconstituted back into an octamer. More recently, however, researchers have turned to the use of recombinant Xenopus laevis histones expressed in E. coli. Not only does this avoid the problem of histone variants and post translational modifications creating a heterogeneous mix of molecules, but it allows for custom tailoring of the histone sequence. In addition to examining histone mutations and adding His-tags, sequence modification can be used to attach reporter groups, DNA damaging agents, or other exotic molecules.

A DNA sequence that assembles into a nucleosome in a predictable translational and rotational manner is known as a nucleosome positioning sequence (NPS). These sequences can occur naturally, or they can be designed synthetically. The most commonly used naturally occurring DNA sequence in nucleosome BER studies is the 5S rDNA sequence, either from Xenopus borealis or Lytechinus variegatus. These sequences, however, can adapt multiple translational positions, in which some
populations of nucleosomes many be translationally shifted in 10 bp increments. Potential multiple translational positions must then be accounted for in experimental design and analysis. The *X. borealis* sequence has been demonstrated to adapt up to five translational settings, while the *L. variegatus* sequence has a least two settings when reconstituted under the nucleosome assembly conditions published by the Luger group (Luger et al., 1999). Heat treatment fixes these nucleosomes to one position, but only for shorter lengths (146 bp). The synthetically derived NPS developed by the Widom lab, named simply “601”, does not have this issue. It adapts one translational position, it is thermodynamically more stable than the 5S sequence, and it is more resistant to destabilization by salt or dilution (Gansen et al., 2009). As such, it has been the preferred substrate in recent years, particularly in generating crystal structures. However, because it is a synthetic substrate, the question arises as to how well the behavior of this sequence reflects the *in vivo* situation.

With regards to studies involving nucleosomes containing linker DNA, it is worth noting that the process of observing histone H1 effect is not a trivial matter, and involves more than simply titrating H1 in with a reconstituted nucleosome. Nonspecific interactions between linker histone and DNA tend to create large aggregates at physiological salt concentrations. *In vivo*, it’s likely that H1 deposition requires a histone chaperone. For example, nucleosome assembly protein-1 (NAP-1) is a histone chaperone for H2A & H2B in mammals, and can also correctly place linker histone on the
nucleosome at physiological salt, without aggregation (Shintomi et al., 2005). Thus, NAP-1 can be used for nucleosome in vitro studies involving linker histone.

When interpreting the results from the following in vitro experiments, context is important. Several groups have estimated the rate of BER in vivo to be ~1 hr. One group characterized depurination levels in IMR90 cells by AP sites biotinylated with an aldehyde-reactive probe. Treatment with H₂O₂ revealed that AP site levels in young IMR90 cells (isolated from 2-4 month old rats) returned to base levels within 45 min (Atamna et al., 2000). A later studied measured levels of four different oxidative lesions via liquid or gas chromatography coupled to mass spec. H₂O₂-treated non-malignant cell lines, AG11134 and HCC1937BL, found repair completed within 1 hr (Nyaga et al., 2007). Thus, a rough time limit is established for asking whether the in vitro BER studies in nucleosomes performed in our lab and others can account for repair in vivo.

1.4.2. Effects of Lesion Orientation and Translational Position

Studies of BER on nucleosomes, performed primarily with glycosylases, have elucidated three general “rules” for repair of a lesion-containing nucleosome. The first rule deals with lesion orientation. Each base within a DNA sequence exists within a preferred helical orientation with respect to the histone octamer (Section 1.1.2). The first rule states that, at least for the first three steps of BER, the rate at which a lesion is processed correlates with whether the damage is oriented outward, away from the
octamer, or inward, buried towards the octamer face. Outward facing lesions are usually, but not always, processed with higher efficiency than inward facing lesions.

The second rule applies to inward facing lesions as well as outward facing lesions that may be occluded in other ways. It states that occluded lesions can be accessed by BER enzymes during periods of spontaneous, transient unwrapping of the nucleosome (see Section 1.1.5). This concept is supported by the fact that increasing enzyme concentration can, to some extent, compensate for lesion occlusion, by increasing the chances to bind the lesion during an unwrapping event.

The third rule relates to translational positioning, that is, how far away a lesion is from the edge of a nucleosome. It states that lesions located closer to the dyad axis will be processed less efficiently than those positioned closer to the edge of the nucleosome. This relates to the second rule in that DNA near the dyad unwraps with far less frequency than DNA near the edge. The studies that led us to make these rules are described below.

An early study by Nilsen et al. examined the human DNA glycosylases UDG and SMUG1 on U:A located 22, 51, or 54 nucleotides away from the dyad axis (Nilsen et al., 2002). They found a ~3-9-fold reduction in glycosylase activity on the nucleosome. However, this early study found no correlation between efficiency and either rotational orientation or translational position. One year later, the Smerdon group examined UDG on nucleosomes containing U:G located at two positions located 4 bp from the dyad in
either direction. They found a similar ~10-fold suppression of activity with respect to naked DNA. However, they found a 2-3-fold rate difference between lesions with opposite helical orientations, with the more outward facing lesion being processed more efficiently.

In 2007, the Pederson lab examined NTHL1 activity on Tg-containing nucleosomes, varying both rotational orientation and translational position (Prasad, Wallace, et al., 2007). NTHL1 cleaved an outward facing Tg at position -51 with efficiency nearing that of naked DNA, whereas an inward facing Tg located only 5 bp away at position -46 was inhibited ~2-3-fold, in agreement with the earlier Smerdon study. An inward facing lesion placed 20 bp further into the nucleosome, at position -26, was inhibited by over 5-fold, but could still be recognized. Importantly, increasing the concentration of NTHL1 resulted in an increase in Tg excision. This was early evidence that the nucleosome unwrapping demonstrated by Widom and others could facilitate access by a DNA glycosylase to an otherwise sterically occluded damage. A second Pederson lab study examined the ability of NEIL1 to excise Tg lesions from nucleosomes (Odell et al., 2010), although this study was complicated by NEIL1’s binding affinity for non-specific DNA. After adjusting for this non-specific binding, they found that NEIL1 can indeed excise Tg from a nucleosome substrate. However, under physiological conditions, it seems unlikely that NEIL1 would be able to locate and process lesions in chromatin with great efficiency, due to this non-specific binding property. This is in line with its proposed in vivo function(s) as a replicative or mitochondrial glycosylase,
wherein lesion search does not involve nucleosomes and occurs over a smaller area. Importantly, this study also experimentally determined the intercellular concentrations of both NEIL1 and NTHL1 to be between 250 and 800 nM. This gives an upper limit on how much glycosylase can reasonably be added to a nucleosome reaction and remain potentially physiologically relevant.

The Hayes group systematically examined the activity of *E. coli* UDG on U:A at eight different translational positions, located from the edge all the way to the dyad (Cole et al., 2010). Three uracils were inward facing, while five were oriented outward. Under enzyme-excess conditions, all 5 outward facing lesions could be processed with only a few fold reduction in efficiency, even as far in as the dyad. However, for inward facing lesions, they observed a sharp drop in efficiency as the uracil was moved closer to the nucleosome center, with a >1000-fold slower rate near the dyad. This study also noted that cleavage of inward facing uracils increased with higher UDG concentrations.

The Smerdon group has also looked at the effect of crosslinking lesion-containing nucleosomes (Hinz et al., 2010). Formaldehyde crosslinking “freezes” histone-DNA contacts without reacting with DNA or perturbing nucleosome structure. This results in reduced rotational flexibility and nucleosome unwrapping. They first examined *E. coli* UDG on three uracils located near the dyad: one inward, one outward, and one in between inward and outward (called “middle”). The lesion orientation rule held true for all three, with the outward lesion processed efficiency as naked DNA. Increasing UDG
concentration from 2 nM to 20 nM approximately doubled processing of the inward and middle uracils. Nucleosome crosslinking further inhibited processing of the inward facing lesion by >2-fold, but had no effect on the middle lesion. The crosslinked outward lesion was actually removed with greater efficiency than naked DNA, which they speculate might be due to crosslinking “locking” the base into a more accessible confirmation. A later study examined six different translational positions, ranging from the dyad to 49 bp away. Again, crosslinking in these experiments reduced inward uracil processing but not outward, ranging from a 2-fold decrease at position -49 to complete inhibition at positions +25 and +4 (Rodriguez & Smerdon, 2013).

The Stivers lab performed a “global assessment” of uracil positions with the Widom 601 sequence (Ye et al., 2012). To generate a population of substrates in which every thymine was converted to uracil, they performed primer extension on 601 template strands with a specific ratio of dUTP to dTTP designed to yield <1 uracil per strand. This allowed them to examine every uracil position within the 601 sequence with only two UDG reactions: one in which the top strand was used as a template, the other using the bottom strand. Analyzing a total of 52 uracil sites, they found that while many of the outward facing sites were processed effectively, there was also population of the outward facing uracils that were unreactive with UDG. These poorly reactive lesions could largely be explained by other physical impediments: these sites had narrow minor grooves, histone-DNA contacts, possible uracil-histone hydrogen bonding, and/or a histone side chain jutting into the minor groove. So while helical orientation is important for lesion
recognition, it is not the only deciding factor. Conversely, UDG demonstrated significant activity on two inward-facing uracils. The first site could likely be explained by its close proximity to the nucleosome edge. The second site, however, was located at position -32 in the top strand and was processed with ~40% the efficiency of naked DNA. This particular site had a wide minor groove, although this was not sufficient to explain the results, as several other occluded uracils had similar groove widths, yet displayed little to no uracil excision. Thus, some other factor(s) are likely involved.

Recently, the Pederson group quantitated nucleosome unwrapping rates during the processing of oxidized lesions (Maher et al., 2013). They placed inward facing thymine glycols at translational positions -26 & -46 in 5S rDNA, and at positions -27 & -47 in 601 substrates. Kinetics were analyzed after adding NTHL1 to each substrate at 2, 15, and 75 nM enzyme (well within in vivo concentrations). Table 3 compares the rates generated by this study to the rates generated by (Tims et al., 2011). The calculated unwrapping rates were fairly consistent with the results from the Widom group, particularly considering the different proteins, binding sites, and methods used. With these data, they concluded that a lesion located at position -46 (5S) or -47 (601) would become transiently exposed approximately once every 20 or 40 minutes, respectively. While this timeframe is within in vivo rates (~1hr), these lesions are only ~1/3 into the nucleosome. The rates for lesions positioned significantly farther in, such as positions -26 and -27, are too low for this timeframe, and are therefore unlikely to be removed in vivo until replication/transcription occurs, or with the aid of chromatin remodelers.
1.4.3. Pol β and Other BER Enzymes on Nucleosomes.

Many of the studies listed above included APE1 with UDG in nucleosome reactions. At high enough concentrations of each, any lesion position could be converted to a strand break, demonstrating that the first two steps of BER, at least, can occur on nucleosomes (Nilsen et al., 2002) (Beard et al., 2003) (Hinz et al., 2010) (Rodriguez & Smerdon, 2013). DNA glycosylases and APE1 are structurally similar, binding only ~1/2 of the DNA helix, and bending their bound DNA substrate 45° to 70°. Consequently, it’s unsurprising that they display similar activities on nucleosomes.

Pol β tends to fare less well on nucleosomes. Although pol β structure is similar to DNA glycosylases and APE1 in that it covers only ~1/2 the DNA helix, it bends DNA at a bound gap by 90° and makes DNA contacts up to 6 bp away on either side of the gap (Pelletier et al., 1996). This may explain some of the comparatively reduced rates of pol β observed on nucleosomes.

The early UDG/SMUG1 study by Nilsen et al. also added pol β to their reactions and found activity at all three positions tested (Nilsen et al., 2002). However, the Smerdon group, while examining lesions near the dyad, added 10 nM pol β to UDG/APE1-cleaved substrates and found no gap filling activity even after 4 hours of incubation (Beard et al., 2003). In a later study, they incubated UDG/APE1-processed nucleosomes with pol β at multiple positions, and found that gaps could be processed, but
at a reduced rate (Rodriguez & Smerdon, 2013). However, even outward-facing locations were inefficiently processed. In this case, it appeared that translational position mattered more than rotational, as pol β experienced a 3-fold greater inhibition of pol β at outward position +10 than at inward position -35, whereas outward +10 and inward +4 were inhibited by a similar 2.5-fold and 3-fold, respectively. Formaldehyde crosslinking performed in this study had a significant inhibitory effect only on the inward facing +4 substrate, indicating that transient unwrapping is likely necessary to access to inward-facing lesions. Pol β also incorporated more nucleotides at position -35 (5 nt) than at positions +4/+10 (1 nt). APE1 addition did not stimulate pol β on activity nucleosomes, although there are conflicting reports as to whether or not APE1 stimulates pol β in naked DNA to begin with (Liu et al., 2007) (Wong & Demple, 2004). It remains possible that APE1 might stimulate pol β in the presence of a 5′-dRP site, but this remains untested in nucleosomes.

The Pederson lab reconstituted the entire BER pathway (NTHL1, APE1, pol β, and LigIIα-XRCC1) on nucleosomes using both outward and inward facing damages located ~1/3 into the nucleosome (Odell et al., 2011). In these experiments, they observed that 3 nM pol β efficiently filled an outward oriented gap, while inward gap filling was suppressed ~3-fold. While this activity was comparatively robust to some of the other studies, pol β was inhibited by an inward facing damage more than the first two steps (~2-fold for NTHL1 and <2-fold for APE1). Remarkably, addition of LigIIα-XRCC1 to pol β enhanced gap filling on the inward site by ~2-fold. This stimulation was also
present in pol β mutants that cannot physically interact with XRCC1, suggesting that the enhancement is due to disruption of histone-DNA contacts by LigIIIα-XRCC1 and not through protein-protein interactions.

There is evidence that the final step of BER, ligation by either LigIIIα or Lig I, can occur on nucleosomes. This is somewhat surprising, as DNA ligases must completely encircle the DNA helix in order to seal a nick. The Hayes lab observed Lig I activity on nucleosomes at position -12, albeit with a 10-fold reduced activity (Chafin et al., 2000). The original UDG/SMUG1 study by Nilsen et al. observed fairly robust ligation by LigIII at outward facing position -51 (Nilsen et al., 2002). Odell et al. found that 5 nM LigIIIα-XRCC1 could ligate nicks in nucleosomes with ~2-3-fold reduced activity (Odell et al., 2011). Notably, they observed no difference between the outward position -51 and the inward position -46. This is an expected result from an enzyme that must fully encircle its DNA substrate. Ligation in nucleosomes is discussed at greater length in Chapter 3.

1.4.4. Linker Studies

Might factors such as linker length or histone H1 placement affect rates of BER in nucleosomes? In the previously described Hayes study examining E. coli UDG activity on U:A at various translational positions, the effect of linker histone H1 on a 227 bp nucleosome substrate was also tested (Cole et al., 2010). They found that the presence of H1 inhibited UDG activity on uracils, but only at certain translational positions located near the dyad or DNA entry/exit points. As well, processing of uracils located within
linker DNA were inhibited by H1. This study did not use NAP-1 to deposit histone H1. Instead, an optimum H1 to nucleosome ratio was determined experimentally, and H1-bound nucleosomes were subsequently purified from agarose gels.

The addition of non-histone associated linker DNA itself can yield striking differences. In a Lig I nucleosome study, the Hayes group compared 154 and 218 bp X. borealis 5S rDNA nucleosomes containing a nick at position -12 (Chafin et al., 2000). Lig I could not ligate the nick in the 154 bp substrate, demonstrating a $10^5$-fold reduced efficiency compared to naked DNA. However, a 218 bp substrate could be ligated with only a 10-fold decrease in activity with respect to naked DNA, a 10,000-fold higher rate than with the 154 bp substrate. Conversely, linker histone H1 had no effect on Lig I activity, although the study gave no information as to how H1-nucleosomes were generated.

Though one might make the assumption that linker DNA located between nucleosomes would be equally as accessible to the BER machinery as DNA flanking a nucleosome end, it is not a foregone conclusion. The Angelov lab examined the accessibility of an 8-oxoG lesion located within linker DNA in a dinucleosome construct containing the 601 sequence (Menoni et al., 2012). Both linker DNA lengths tested (20 and 75 bp) demonstrated 8-oxoG excision rates by OGG1 to be similar to naked DNA, confirming that linker DNA between nucleosomes remains enzymatically accessible. However, deposition of linker histone H1 onto nucleosomes with NAP-1 reduced OGG1
efficiency ~10-fold. Subsequent removal of H1 by higher concentrations of NAP-1 restored the ability of OGG1 to remove 8-oxoG. Addition of ATP-dependent chromatin remodeler RSC was insufficient to enhance OGG1 activity when H1 was present. This suggests that linker DNA in vivo is likely occluded from BER by bound linker histone H1; additionally, the presence of H1 is resistant to chromatin remodeling by RSC, but lesion accessibility could instead be facilitated by high concentrations of NAP-1.

1.4.5. Histone PTMs and Nucleosome BER

Post-translational acetylation of histones tends to increase DNA unwrapping events. Limited trypsin digests of the histone tails is one method used to mimic nucleosome hyperacetylation. The Smerdon group tested this with human UDG, APE1 and, pol β, and found that histone tail removal from chicken erythrocyte core particles did not affect efficiency of any of the three enzymes at positions located 4 nt from either side of the dyad (Beard et al., 2003).

In contrast, the Hayes group found that Lig I activity on nucleosomes containing a nick at position -12 benefited significantly from tail removal (Chafin et al., 2000). However, Sac I cleavage of a site located close to the nick position was unaffected by the absence of histone tails. This meant that the increased ligation efficiency was not due to an increased frequency of nucleosome unwrapping events. Instead, they speculate that the histone tails were making contacts with the 154 bp substrate that could somehow be
disrupted by Lig I in a 218 bp substrate (see linker section above), but not in the shorter DNA.

Ishibashi et al. examined human DNA glycosylase MBD4 activity on nucleosomes containing outward-facing T:G mismatches located 14 and 44 nucleotides from the dyad (Ishibashi et al., 2008). MBD4 could excise thymine from both positions, but with a 2-3-fold reduction compared to naked DNA. However, after hyperacetylation (10-12 acetyl groups per octamer) of nucleosomes, excision rates jumped back to nearly that of naked DNA, indicating that acetylation stimulates MBD4 activity on nucleosomes.

Most recently, the Smerdon group constructed both 5S rDNA and 601 nucleosomes with specifically acetylated histone H3 at either Lys14 or Lys56 (Rodriguez et al., 2016). Acetylation of Lys14, which is located on the N-terminal tail, causes the ATP-dependent chromatin remodeler RSC to bind (Kasten et al., 2004). Additionally, acetylation at Lys14 tags nucleosomes for eviction from promoters by histone chaperone NAP-1 (Luebben et al., 2010). Lys56 is located in the histone fold near the DNA entry/exit region; acetylation of Lys54 reduces nucleosome stability. Acetylated Lys54 is found in newly synthesized H3, but also persists when DSBs are present (Masumoto et al., 2005). Although acetylated Lys56 disrupts nucleosome contacts, restriction digest accessibility was unaffected in either Lys14 or Lys56. The group examined inward and outward positions in both 5S and 601, as well as multiple translational positions in 601.
They observed no difference in activity of UDG on uracils in either of the acetylated nucleosomes at any position. However, surprisingly, both H3 acetylations inhibited pol β gap-filling by 2-5-fold.

1.4.6. Chromatin Remodelers and Nucleosome BER

The Angelov group engineered 601 nucleosomes with an 8-oxoG lesion located 10 bp from the dyad (a “middle” facing site) to examine the activity of mouse OGG1 on nucleosomes (Menoni et al., 2007). Similar to studies with other glycosylases, they observed at ~10-fold inhibition on the nucleosome substrate using 0.25 U OGG1, although at 2 U OGG1, inhibition drops to ~3-fold. From here, they compared normal nucleosomes with nucleosomes containing the variant histone H2A.Bbd. H2A.Bbd nucleosomes are located in transcriptionally active DNA and have a more relaxed, less stable structure, but are resistant to remodeling by SWI/SNF (Doyen et al., 2006) (Angelov et al., 2004). Here, OGG1 was still inhibited by H2A.Bbd, but to a lesser degree than normal nucleosomes.

Next, they repeated the reactions but included 0.3 to 1.2 U of SWI/SNF. At lower concentrations, SWI/SNF catalyzes nucleosome mobilization, sliding the octamer to a different position. At higher concentrations (>1 U), SWI/SNF remolds (disrupts) nucleosomes. As H2A.Bbd nucleosomes are resistant to remodeling, addition of SWI/SNF did not increase excision of 8-oxoG by OGG1. In mobilized H2A nucleosomes, however, OGG1 demonstrated 25-30% higher activity, and reached nearly
naked DNA levels in the remodeled H2A nucleosomes. Interestingly, SWI/SNF stimulates the NER pathway only 1.5 to 2-fold on nucleosomes in vitro (Hara & Sancar, 2003) (Hara & Sancar, 2002).

In an effort to study BER on a 30 nm fiber, a study from the Smerdon group reconstituted 12 tandem repeats of a 5S rDNA plasmid into uracil-containing nucleosomes (Nakanishi et al., 2007). Uracils were generated by deamination of cytosines with sodium bisulfite, while salt-induced higher order structure formation was confirmed by sedimentation velocity analysis. These arrays were then incubated with 1 or 10 nM UDG and APE1 and compared to naked DNA controls. At 1 nM, repair occurred at a ~2-3-fold slower rate than naked DNA, whereas at 10 nM, the arrays were digested to completion. At the end of the 10 nM UDG/APE1 digest, pol β was added to 1 nM and was observed to be able to incorporate dCTP into the cleaved sites at ~5-fold reduced efficiency. Adding pol β at 10 nM reduced inhibition to <3-fold. This demonstrated that the first three steps of BER could occur within a higher order chromatin structure, albeit at a reduced rate. ISW1 and ISW2 are chromatin remodelers that can catalyze nucleosome sliding on a nicked DNA template, making them a good choice for examining possible stimulation of pol β by chromatin remodeling. Following digest with UDG/APE1, either ISW1 or IWSI2 was added to the reactions at a 1 to 1 ratio with nucleosomes, and the pol β experiment was repeated. ISW1 and ISW2 stimulated gap filling by pol β 4- and 6-fold, respectively.
1.5. Figures

Figure 1-1. Nucleosome structure & histone fold domains.

Figure from (Cutter & Hayes, 2015).

(A) Model of a nucleosome core (PDB: 1KX5 (Davey et al., 2002)). (B) Schematic of the four core histone proteins, indicted by colored text. The histone fold consists of α1-3 plus two loops located between the helices, and is depicted in (C). All four histones contain this fold as well as disordered N-terminal tails of various lengths. H2A additionally contains a C-terminal tail and “docking domain” which forms contacts with the H3-H4 tetramer. (D) Architecture of the nucleosome. Approximately one half has been removed for enhanced clarity. Four helix bundles are highlighted by black ovals.
Figure 1-2. Differing models for the 30 nm fiber

Figure from (Khorasanizadeh, 2004).
See text (page 10) for description.
In the simplest scenario, leading strand synthesis is halted by a single-strand break (SSB). Collapse of the replication fork converts the SSB into a one-sided DSB. Attempted repair by NHEJ would potentially lead to chromosomal rearrangements or translocations. Several kinds of DNA lesions, including thymine dimers, certain oxidized bases (e.g. thymine glycol), abasic sites, and inter-strand crosslinks (such as those caused by cisplatin) can cause replication forks to stall. The replication fork can stall and regress in an attempt to repair the blocking damage. Fork regression will cause newly synthesized leading and lagging strands to reanneal, as depicted. The newly synthesized lagging strand can serve as a template to extend the leading strand, producing the “chicken foot” intermediate shown. Cleavage of the chicken foot by resolvases, will produce, once again, a one-sided DSB.

---

**Figure 1-3. DSB formation via replication fork collapse.**
Figure 1-4. Various products of endogenous reactive oxygen species or UV radiation.

Figure from (Slupphaug et al., 2003).
Figure 1-5. MMR in eukaryotes

Figure adapted from (Fukui, 2010). See text for description.
Figure 1-6. NER in eukaryotes

Figure from (Lans et al., 2012). See text for description.
Figure 1-7. DSBR pathways

Figure from adapted from (Iliakis et al., 2015). See text for description.
Figure 1-8. BER and subpathways

Figure from (Krokan & Bjoras, 2013). See text for description.
Figure 1-9. DNA glycosylase catalysis, mono- and bi-functional

Figure from (Fromme et al., 2004). See text for description.
Figure 1-10. AP endonuclease 1.

Figure adapted from (Li & Wilson, 2014).
Top: Schematic of APE1. Bottom left: crystal structure of APE1 bound to un-incised DNA, with recognition loops indicated and the abasic site circled (PDB: 1DE8 (Mol et al., 2000)). Note the ~35° bending of the DNA. Bottom right: APE1 bound to incised DNA (PDB: 1DE9 (Mol et al., 2000)). Active site residues are E96, D210, and H309.
Figure 1-11. Ligase mechanism of action.

Figure from (Wikimedia Commons). See text for description.
Figure 1-12. LigIIIα-XRCC1

Figure from (Ellenberger & Tomkinson, 2008).
The above schematic details the major domains of XRCC1 and LigIIIα, as well as major interaction partners. The regions between XRCC1’s NTD-BRCTI and BRCTI-BRCTII are intrinsically disordered, allowing for multiple interactions with other repair and accessory proteins. Some interactions, such as PNK, APTX, and APLF with XRCC1, are phosphorylation-dependent (not depicted in figure).
Table 1-1. Formation and deactivation of reactive oxygen species (ROS)

<table>
<thead>
<tr>
<th>Generation of ROS</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\text{H}_2\text{O} \rightarrow \text{H}_2\text{O}^+ + e^-$</td>
</tr>
<tr>
<td>$\text{H}_2\text{O}^+ + \text{H}_2\text{O} \rightarrow \text{OH}^\cdot + \text{H}_3\text{O}$ (hydroxyl radical)</td>
</tr>
<tr>
<td>$\text{OH}^\cdot + \text{OH}^\cdot \rightarrow \text{H}_2\text{O}_2$ (hydrogen peroxide)</td>
</tr>
<tr>
<td>$e^{aq} + \text{O}_2 \rightarrow \text{O}_2^\cdot$ (superoxide radical)</td>
</tr>
<tr>
<td>$2\text{O}_2^\cdot + 2\text{H}^+ \rightarrow \text{O}_2 + \text{H}_2\text{O}_2$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Protection against ROS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Superoxide dismutases (cytoplasm, mitochondria, extracellularly)</td>
</tr>
<tr>
<td>$2\text{O}_2^\cdot + 2\text{H}^+ \rightarrow \text{H}_2\text{O}_2 + \text{O}_2$</td>
</tr>
<tr>
<td>Catalase removes hydrogen peroxide</td>
</tr>
<tr>
<td>$2 \text{H}_2\text{O}_2 \rightarrow \text{H}_2\text{O} + \text{O}_2$</td>
</tr>
</tbody>
</table>

Table adapted from (Slupphaug et al., 2003).
Table 1-2. Mammalian DNA Glycosylases

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>Subcellular localization</th>
<th>Mono-/bifunct.</th>
<th>Substrates &amp; (minor substrates)</th>
<th>Mouse knockout</th>
<th>Human diseasea</th>
</tr>
</thead>
<tbody>
<tr>
<td>UNG2</td>
<td>Nuc</td>
<td>M</td>
<td>U, 5-FU in ss and dsDNA, U:A and U:G context (alloxan, 5-hydroxyuracil, isodialuric acid)</td>
<td>Partial defect in CSR, skewed SHM, B-cell lymphomas.</td>
<td>Complete defect in CSR, HIGM syndrome, infections, lymphoid hyperplasia</td>
</tr>
<tr>
<td>UNG1</td>
<td>Mito</td>
<td>M</td>
<td>Like UNG2</td>
<td>Unknown</td>
<td>Unknown</td>
</tr>
<tr>
<td>SMUG1</td>
<td>Nuc</td>
<td>M</td>
<td>5-hmU, U:G &gt; U:A &gt; ssU, 5-FU, εC in ss and dsDNA</td>
<td>Viable and fertile, SMUG1/UNG/MSH triple k.o. reduced longevity.</td>
<td>Unknown</td>
</tr>
<tr>
<td>TDG</td>
<td>Nuc</td>
<td>M</td>
<td>U:G &gt; T:G (5-hmU in dsDNA, 5-FU)</td>
<td>Embryonic lethal, epigenetic role in development.</td>
<td>Unknown</td>
</tr>
<tr>
<td>MBD4   (MED1)</td>
<td>Nuc</td>
<td>M</td>
<td>U:G and T:G, 5-hmU in CpG context (εC, 5-FU in dsDNA)</td>
<td>Viable and fertile, C to T transitions, intestinal neoplasia.</td>
<td>Mutated in carcinomas with microsatellite instability</td>
</tr>
<tr>
<td>MPG    (AAG)</td>
<td>Nuc</td>
<td>M</td>
<td>3meA, 7meG, 3meG, Hx, εA</td>
<td>Viable and fertile, triple knockouts in MPG/AlkBH2/AlkBH3 hypersensitive to inflammatory bowel disease.</td>
<td>Unknown</td>
</tr>
<tr>
<td>OGG1</td>
<td>Nuc</td>
<td>M/B</td>
<td>8-oxoG:C, Fapy:C</td>
<td>Viable and fertile, OGG1/MUTYH double knockouts cancer prone.</td>
<td>OGG1 activity associated with CAG repeat expansion in Huntington’s disease</td>
</tr>
<tr>
<td>MYH</td>
<td>Nuc</td>
<td>M</td>
<td>A opposite 8-oxoG/C/G</td>
<td>OGG1/MUTYH double knockouts cancer prone</td>
<td>MUTYH variants associated with colon polyposis</td>
</tr>
<tr>
<td>NTHL1</td>
<td>Nuc</td>
<td>B</td>
<td>Tg, FapyG, 5-hC, 5-hU in dsDNA</td>
<td>Viable and fertile, NTHL1/NEIL1 double knockouts cancer prone. Telomeric defects.2</td>
<td>Germline nonsense mutation linked to adenomatous polyposis. Several polymorphisms found in lung cancer patients.4</td>
</tr>
<tr>
<td>NEIL1</td>
<td>Nuc</td>
<td>B</td>
<td>Tg, FapyG, FapyA, 8-oxoG, 5-hU, DHU, Sp and Gh in ss and dsDNA; Sp and Gh in quadruplex DNA1</td>
<td>Viable and normal at birth, obese after 7 months, NTHL1/NEIL1 double knockouts cancer prone.</td>
<td>Unknown</td>
</tr>
<tr>
<td>Enzyme</td>
<td>Subcellular localization</td>
<td>Mono-/ bifunct.</td>
<td>Substrates &amp; (minor substrates)</td>
<td>Mouse knockout</td>
<td>Human disease&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>--------</td>
<td>--------------------------</td>
<td>----------------</td>
<td>--------------------------------</td>
<td>----------------</td>
<td>--------------------------</td>
</tr>
<tr>
<td>NEIL2</td>
<td>Nuc</td>
<td>B</td>
<td>Similar to NEIL1, but does not repair quadruplex DNA</td>
<td>Viable and fertile, older mice accumulate damage, esp. within transcribed regions.&lt;sup&gt;3&lt;/sup&gt;</td>
<td>Unknown</td>
</tr>
<tr>
<td>NEIL3</td>
<td>Nuc</td>
<td>M/B</td>
<td>FapyG, FapyA, Sp and Gh in ssDNA; Sp, Gh, and Tg in quadruplex DNA&lt;sup&gt;1&lt;/sup&gt;</td>
<td>Viable and fertile, memory and learning deficit.</td>
<td>Unknown</td>
</tr>
</tbody>
</table>

<sup>a</sup>Association of germline alterations in DNA glycosylases NTHL1, NEIL1, NEIL2, MPG, TDG, UNG, and SMUG1 with colorectal cancer has been reported, but significance uncertain (Broderick et al., 2006). bifunct., bifunctional; k.o., knockout.

Table modified slightly from (Krokan & Bjoras, 2013). Red text indicates updated findings.

<sup>1</sup>(Zhou, Liu, et al., 2013) (Zhou, Fleming, et al., 2015)  
<sup>2</sup>(Vallabhaneni et al., 2013)  
<sup>3</sup>(Chakraborty et al., 2015)  
<sup>4</sup>(Weren et al., 2015) (Couto et al., 2015)
Table 1-3. Nucleosome unwrapping rates.

<table>
<thead>
<tr>
<th></th>
<th>601</th>
<th>601</th>
<th>601</th>
<th>5S (+1)</th>
<th>5S (0)</th>
<th>601</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Protein probe</strong></td>
<td>LexA</td>
<td>LexA</td>
<td>LexA</td>
<td>hNTH1</td>
<td>hNTH1</td>
<td>hNTH1</td>
</tr>
<tr>
<td><strong>Distance of target sequence from the nucleosome edge (bp)</strong></td>
<td>17.5</td>
<td>27.5</td>
<td>37.5</td>
<td>16</td>
<td>26</td>
<td>27</td>
</tr>
<tr>
<td><strong>κ (s⁻¹)</strong></td>
<td>4.1 ± 0.2</td>
<td>0.016 ± 0.003</td>
<td>0.0017 ± 0.0010</td>
<td>0.13 ± 0.03</td>
<td>0.003 ± 0.0009</td>
<td>0.0014 ± 0.0004</td>
</tr>
<tr>
<td><strong>Average lifetime of occluded state(s): τ (s)</strong></td>
<td>0.24</td>
<td>63</td>
<td>590</td>
<td>7.6</td>
<td>330</td>
<td>710</td>
</tr>
</tbody>
</table>

Table adapted from (Maher et al., 2013) and (Tims et al., 2011). See text for description.
CHAPTER 2: NUCLEOSOMES SUPPRESS THE FORMATION OF DOUBLE-STRANDED DNA BREAKS DURING ATTEMPTED BASE EXCISION REPAIR OF CLUSTERED OXIDATIVE DAMAGES

Wendy J. Cannan, Betty P. Tsang, Susan S. Wallace and David S. Pederson

From the Department of Microbiology and Molecular Genetics, University of Vermont, Stafford Building, Room 220, 95 Carrigan Drive, Burlington, VT, 05405, USA, Tel.: (802) 656-8587; E-mail: david.pederson@uvm.edu

2.1 Abstract

Exposure to ionizing radiation can produce multiple, clustered oxidative lesions in DNA. The near simultaneous excision of nearby lesions in opposing DNA strands by the Base Excision Repair (BER) enzymes can produce double-strand DNA breaks (DSBs). This attempted BER accounts for many of the potentially lethal or mutagenic DSBs that occur in vivo. To assess the impact of nucleosomes on the frequency and pattern of BER-dependent DSB formation, we incubated nucleosomes containing oxidative damages in opposing DNA strands with selected DNA glycosylases and human apurinic/apyrimidinic endonuclease 1. Overall, nucleosomes substantially suppressed DSB formation. However, the degree of suppression varied as a function of [i] the lesion type and DNA glycosylase tested; [ii] local sequence context and the stagger between opposing strand lesions; [iii] the helical orientation of oxidative lesions relative to the underlying histone
octamer; and [iv] the distance between the lesion cluster and the nucleosome edge. In some instances, the binding of a BER factor to one nucleosomal lesion appeared to facilitate binding to the opposing strand lesion. DSB formation did not invariably lead to nucleosome dissolution and, in some cases, free DNA ends resulting from DSB formation remained associated with the histone octamer. These observations explain how specific structural and dynamic properties of nucleosomes contribute to the suppression of BER-generated DSBs. These studies also suggest that most BER-generated DSBs will occur in linker DNA and in genomic regions associated with elevated rates of nucleosome turnover or remodeling.

2.2 Introduction

Ordinary oxidative metabolism produces hydroxyl radicals and other reactive oxygen species (ROS) that, in large part, are neutralized by antioxidant molecules and enzymes. However, enough ROS remain to generate some ~10,000 oxidative base lesions, and a similar number of abasic sites, in the DNA of every nucleated human cell, every day (Friedberg et al., 2005). These lesions tend to be widely distributed, and most are repaired in an error-free fashion, via base excision repair (BER). In its simplest “short-patch” form, BER begins with the excision of oxidatively damaged bases by DNA glycosylases, or with the excision of ROS-generated abasic sites by human apurinic/apyrimidinic endonuclease 1 (hAPE1). The result in either case is a single-strand DNA gap, which is filled by DNA polymerase β, and sealed by DNA ligase IIIα [for reviews, see: (David et al., 2007; Duclos et al., 2012; Fromme et al., 2004; Hazra et al.,
Ionizing radiation consists of high energy photons that can strip electrons from atoms (Einstein, 1905). A single high-energy photon can disrupt multiple water molecules, thereby generating multiple hydroxyl radicals. When this occurs in the vicinity of DNA, each of the hydroxyl radicals formed can generate an oxidative lesion in DNA. The half-life of hydroxyl radicals in aqueous solutions is about one nanosecond, giving them a reactive radius of ~20Å [reviewed in (Halliwell, 2001)]. Consequently, lesions generated by a single photon tend to be clustered. Near-simultaneous, attempted BER of closely opposed lesions may produce single strand repair intermediates on opposing strands, resulting in formation of a double-strand DNA break (DSB), as depicted in Fig. 1A. This phenomenon has been demonstrated to occur in vitro with naked DNA substrates [reviewed in (Eccles et al., 2011; Sage & Harrison, 2011)], and multiple lines of evidence indicate that it occurs in vivo as well (Blaisdell et al., 2001; Blaisdell & Wallace, 2001; Yang et al., 2006; Yang et al., 2004). In particular, mutation or reduced expression of DNA glycosylases decreased the frequency of DSBs induced by ionizing radiation, in both prokaryotes and eukaryotes. Over-expression of DNA glycosylases increased the frequency of DSBs in cells exposed to ionizing radiation, but not in cells exposed to oxidizing agents that generate single, isolated DNA lesions.

In eukaryotes, BER and other DNA transactions occur in chromatin, where nucleosomes restrict access to DNA. Both mono- and bi-functional DNA glycosylases can initiate the BER of oxidative lesions in nucleosomes, albeit with varying and
sometimes much reduced efficiency relative to their activities on naked DNA [reviewed in (Odell et al., 2013); see also (Cadet et al., 2012; Maher et al., 2013; Ye et al., 2012)]. The efficiency with which a nucleosomal lesion is processed depends critically on its helical orientation relative to the underlying histone octamer. These observations suggested that the packaging of DNA in nucleosomes would substantially suppress BER-mediated DSB formation in vivo. In particular, only one of two closely spaced lesions on opposing strands could be optimally accessible to DNA glycosylases. In such cases, BER of the more readily accessible of two closely opposed lesions might go to completion before BER of the less accessible lesion commenced, thereby precluding DSB formation. Here, we have tested this prediction, using nucleosomes that contain discretely positioned oxidative lesions in opposing DNA strands. We report that nucleosomes partially protect DNA from BER-mediated DSB formation, but that the protection is not absolute. We have also investigated the enzymatic and structural factors that influence the efficiency with which DSBs occur in nucleosomes. Our results predict that the distribution of BER-mediated DSBs within the genome is non-random, and very likely influenced by cellular processes that entail remodeling or transient disruption of nucleosomes.

2.3 Experimental Procedures

Preparation of nucleosome-length DNA substrates containing oxidative lesions.

We used two methods to assemble mononucleosome-length (170 to 185 bp) DNAs, containing multiple thymine glycol (Tg) or 7,8-dihydro-8-oxoguanine (8-oxoG) lesions. The first began with the synthesis of a full-length single-strand DNA template in
a linear PCR reaction. The template for this reaction was prepared by linearizing the plasmid pBS-5SLv (Prasad, Wallace, et al., 2007) at a Kpn I site immediately adjacent to the 5S ribosomal DNA-containing nucleosome positioning sequence from *Lytechinus variegatus*. To this DNA, we added buffer, Taq DNA polymerase (Invitrogen), and a 25-fold molar excess of a lesion-containing primer (Midland Certified Reagent Co.) that had been 5’ end-labeled with $\gamma$-$^{32}$P-ATP by polynucleotide kinase (New England Biolabs). After 22 cycles of denaturation (95ºC), annealing (53ºC), and extension (72ºC), we purified the resulting “top strand” DNA by denaturing PAGE, quantified it by scintillation counting, and mixed it with equimolar amounts of three “bottom strand” oligodeoxyribonucleotides (oligos). Of the bottom strand oligos, the 5’-most was $\gamma$-$^{32}$P end-labeled, making it possible to individually monitor excision of oxidative lesions from both top and bottom strands of DNA. In later experiments, we prepared nucleosome-length DNA substrates, using the same bottom strand DNA oligos as before, along with three “top strand” oligos. To these, we added buffer and a thermostable DNA ligase (Ampligase, from Epicentre), and a small amount (1/20th) of undamaged template DNA (from pBS-5SLv), to ‘guide’ the initial ligation events. Reactions were allowed to proceed for 297 cycles of denaturation and ligation. The resulting full-length products were purified by denaturing PAGE, annealed to one another, and treated with Exonuclease I (New England Biolabs) to remove any remaining single-stranded DNA.

**Nucleosome reconstitutions.**
In early experiments, the lesion-containing DNAs described above were assembled into nucleosomes by octamer transfer, using a 100 to 200-fold excess of chromatin donor prepared from chicken erythrocytes as described (Prasad, Wallace, et al., 2007). In later experiments, lesion-containing DNAs were mixed with an 80-fold molar excess of unlabeled carrier DNA (isolated from micrococcal nuclease-digested chicken chromatin), and a slight molar excess of purified histone octamers (assembled using recombinant human or Xenopus histones from Escherichia coli (Dyer et al., 2004; Luger et al., 1999)), in 25 mM HEPES [pH 8.0], 1 mM EDTA, 1 mM DTT and 2 M sodium chloride, and dialyzed into the same buffer lacking sodium chloride. In both methods, the lesion-containing nucleosomes were adjusted to 2.5 nM and diluted to 0.5 nM as needed for activity assays. The presence of unlabeled carrier nucleosomes ensured that the total concentration remained well above concentrations at which one sees dilution-driven dissociation of nucleosomes [c.f. (Pederson et al., 1986)]. To assess reconstitution efficiencies, nucleosomes were electrophoresed through 5% native gels in 0.5-X Tris-Borate-EDTA buffer. The impact of small amounts of contaminating naked DNA on experimental results was eliminated computationally, as previously described (Prasad, Wallace, et al., 2007).

**Isolation of BER enzymes.**

Wild type human endonuclease III (hNTH1) and a variant lacking the first 55 amino acids (hNTH1∆55) were expressed and purified as described (Liu & Roy, 2002;
Odell et al., 2010). hAPE1 and *E. coli* formamidopyrimidine-DNA glycosylase (Fpg) were purchased from New England Biolabs.

Human 8-oxoguanine (hOGG1) glycosylase cDNA (isoform 1a) containing a C-terminal 6xHis-tag was cloned into the vector pET30 and transformed into Rosetta 2 *E. coli* cells (Novagen). To auto-induce transcription of the hOGG1 gene, transformants were grown for 20-48 hours in “Terrific Broth” containing 0.5% glycerol, 0.05% glucose, and 0.2% α-lactose monohydrate (Studier, 2005). Cells were harvested and lysed by sonication (4 x 3 minutes) in 50 mM sodium phosphate, pH 8.0, 10 mM imidazole, 150 mM sodium chloride, 10% glycerol and 5 mM β-mercaptoethanol, supplemented with “cOmplete” protease inhibitor cocktail (Roche). The lysates were cleared by centrifugation at 30,000 x g for one hour, and loaded onto a 5 mL nickel chelating column (GE Healthcare), pre-equilibrated in lysis buffer. The column was eluted with a linear 10-500 mM imidazole gradient in lysis buffer. Fractions containing hOGG1 were identified using SDS-PAGE, pooled, dialyzed into 20 mM HEPES, pH 7.6, 150 mM sodium chloride, 10% glycerol and 5 mM β-mercaptoethanol, and loaded onto a 5 mL SP FF sepharose column (GE Healthcare), pre-equilibrated in the same buffer. hOGG1 was eluted with a linear 0.15 – 1 M gradient of sodium chloride in loading buffer. Fractions containing hOGG1 were identified as above, pooled, dialyzed into 20 mM HEPES, pH 7.6, 100 mM sodium chloride, 50% glycerol and 1 mM DTT, and stored at -20°C.

**Enzyme activity assays.**
All the glycosylase enzyme concentrations reported in this study refer to active enzyme, as determined using a Schiff base assay (Verdine & Norman, 2003). The reported concentrations of hAPE1 refer to total enzyme, calculated from data supplied by the manufacturer. Unless otherwise noted, enzyme reactions also contained 0.5 nM lesion-containing nucleosomes or naked DNA in 25 mM HEPES, pH 8.0, 1 mM EDTA, 1 mM DTT, 100 mM sodium chloride, 5 mM MgCl₂, 0.05% NP-40 and 0.1 mg/mL bovine serum albumin (New England Biolabs). Naked DNA control reactions contained either chicken chromatin or micrococcal nuclease-digested chicken DNA (depending upon nucleosome reconstitution method), such that carrier conditions matched those used in the nucleosome reactions. Reactions were carried out at 37°C for up to 30 min, and quenched by the addition of 1/5th volume of 2.5% SDS, 250 mM EDTA, and 2.5 mg/mL of freshly added proteinase K. After a 30-60 min incubation in quench buffer at 37°C, aliquots to be used for quantifying double-strand break (DSB) formation were mixed with equal volumes of 100 mM tris base, 25 mM EDTA, 0.2% SDS, 2 mg/ml bromophenol blue and 12% glycerol, and electrophoresed through 6% native gels. Aliquots to be used for quantifying single-strand break (SSB) formation were mixed with equal volumes of formamide, containing 20 mM EDTA and 2 mg/ml each of bromophenol blue and xylene cyanol, and fractionated on 6% sequencing gels. Gels were then dried, exposed to K-screens (BioRAD), and bands of interest quantified using a PharosFX Plus phosphoimager with Quantity One software (BioRAD). After adjusting for backgrounds (using the global subtraction method), DSBs were calculated as a percentage of total DNA. SSBs were calculated for each strand by dividing the individual product band
Assays for nucleosome integrity during attempted repair of clustered lesions.

To assess the fate of clustered lesion-containing nucleosomes during attempted BER, aliquots were removed from the above-described reactions after 30 minutes at 37°C, mixed with equal volumes of 25 mM HEPES, pH 8.0, and 10% glycerol, and immediately subjected to electrophoresis on a 6% native gel. A portion of each aliquot was reserved for quantification of DSB formation, as described above. In some instances, yet another aliquot was treated with 50 units of Age I, a restriction endonuclease whose target in 5S rDNA nucleosomes is normally inaccessible (Prasad, Wallace, et al., 2007). To assess Age I cleavage efficiencies, samples were processed and fractionated as in the assay of DSB formation.

2.4 Results

Experimental Design.

DNA glycosylases discover oxidative lesions by extending residues into the minor groove of DNA and inducing damaged bases to flip through the major groove into an extra-helical configuration (Bergonzo et al., 2011; Parker et al., 2007; Qi et al., 2009). Steric constraints imposed by the histone octamer or DNA in the adjacent gyre may impede either glycosylase binding or base flipping. As a result, DNA glycosylases can bind directly to only a small fraction of the base lesions that may form in nucleosomes.
[reviewed in (Odell et al., 2013); see also (Cadet et al., 2012; Maher et al., 2013; Ye et al., 2012)]. This is illustrated in Figs. 1B & C, which depict the region of a nucleosome where we inserted base lesions for our initial studies. In naked DNA, attempted BER of these lesions might generate a DSB, whereas in a nucleosome, only the lesion located 51 nt from the dyad axis is oriented so as to permit direct binding of a DNA glycosylase. Binding of a DNA glycosylase to the other, less optimally oriented lesions (located 46 and 49 nt from the dyad axis on the opposite strand) can occur only when they are exposed to solvent during episodes of spontaneous partial unwrapping of DNA from the histone octamer (Maher et al., 2013; Prasad, Wallace, et al., 2007). Because this DNA unwrapping-dependent mode of BER is much less efficient, it seemed likely that nucleosomes inhibit the near-simultaneous lesion processing events that could lead to DSB formation. This reasoning was the basis for a null hypothesis, namely that nucleosomes fully protect DNA from BER-mediated formation of DSBs. In an attempt to disprove this hypothesis, we constructed and tested an array of model nucleosomes, each containing opposing strand base lesions. To fix the helical orientation of the lesions relative to the histone octamer, nucleosomes contained either the naturally-occurring Lytechinus variegatus 5S ribosomal DNA or the synthetic 601 DNA nucleosome positioning sequence. DNA in the 601 nucleosome occupies a single discrete helical and translational position relative to the underlying histone octamer (Lowary & Widom, 1998; Thastrom et al., 1999). In the 5S rDNA nucleosomes, most DNA molecules occupy one of three translational positions, separated by 10 bp intervals (Dong et al., 1990; Flaus et al., 1996; Prasad, Wallace, et al., 2007). While the relative abundance of these
translational variants differs somewhat with reconstitution conditions (Maher et al., 2013), the fact that they are separated by an integral number of helical turns means that the DNA in each nucleosome has the same helical orientation with respect to the underlying histone octamer (Prasad, Wallace, et al., 2007; Simpson & Stafford, 1983).

**Impact of the stagger between opposing lesions on double-strand break formation.**

In nucleosomes, a base lesion that is optimally positioned for processing will be paired with an occluded base in the opposing strand. Thus, it was likely that optimal processing of opposing strand lesions would require that they be offset, or staggered, relative to one another. Two factors guided our decisions on the range of offsets tested. First, the short reactive lifetimes of hydroxyl radicals produced by a single photon suggested that clustered DNA damages generally lie within ~10 bp of one another. This places an upper limit on the distance between opposing strand lesions. To ensure that both lesions would be at least partly accessible to DNA glycosylases in nucleosomes, we adopted an even more stringent upper limit of five bp between opposing strand lesions. Second, efficient DSB formation occurs in naked DNA substrates only when opposing strand lesions are offset by at least three bp (Harrison et al., 1999). This observation is consistent with evidence that most DNA glycosylases interact with both the DNA lesion and two or more backbone residues on either side of the lesion [e.g. (Fromme & Verdine, 2003; Tchou et al., 1993)]. Thus, an AP site or DNA gap located just one or two bp away from a base lesion in the opposing strand would likely reduce the affinity of a DNA glycosylase for its substrate. This consideration led us to construct as a negative control, a
nucleosome with a two bp stagger between opposing strand lesions. We thus replaced thymine residues in each strand of DNA with the oxidative lesion thymine glycol (Tg), generating DNA with opposing lesions separated by either two or five bp ("Tg51-2" and "Tg51-5", respectively). When assembled into nucleosomes, the top strand Tg was located 61, 51, or 41 nt from the central dyad axis, with the majority at position 51 (Prasad, Wallace, et al., 2007). The bottom strand Tg was either two or five nt closer to the dyad axis, as depicted in Fig 1B & C.

hNTH1 is one of two bifunctional human DNA glycosylases that we knew from earlier studies could excise single Tg residues from nucleosomes. As outlined above, we predicted that processing of either lesion in Tg51-2 DNA would degrade the binding site needed to process the opposing strand lesion. However, because of its exceptionally low turnover rate, it was possible that hNTH1 binding to its product might physically prevent binding of a second hNTH1 molecule to an opposing strand lesion. To eliminate this possibility, we probed Tg51-2 nucleosomes and DNA with hNTH1Δ55, an N-terminal truncation mutant that exhibits an elevated turnover rate but is otherwise fully active (Liu & Roy, 2002; Odell et al., 2011). As can be seen in Figs. 2B & D, DSB formation was poor in naked DNA and virtually undetectable in nucleosomes. By contrast, hNTH1Δ55 formed DSBs in a large fraction (72±7%) of the naked Tg51-5 DNA molecules and in a smaller but still significant fraction (21±6%) of Tg51-5 nucleosomes (Figs. 2C & D).

hNTH1-mediated, double-strand break formation in nucleosomes is lyase-limited.
Although the reduced yield of DSBs in Tg51-5 nucleosomes as compared to naked DNA (Fig. 2) could be due entirely to steric factors, it was possible that the results were influenced by other properties of hNTH1. In particular, the lyase activity in hNTH1 is substantially lower than the glycosylase activity; this is also the case for several other bi-functional DNA glycosylases (Marenstein et al., 2003; Parikh et al., 1998; Pope et al., 2005; Pope et al., 2002; Sung & Mosbaugh, 2000; Yang et al., 2001). Human apurinic-apyrimidinic endonuclease 1 (hAPE1) will displace hNTH1 from abasic sites and increase production of SSBs (Prasad, Wallace, et al., 2007). Fig. 3A & B show that addition of hAPE1 increased both the rate and extent of DSB formation in both nucleosomes and DNA, but did not significantly alter the degree to which nucleosomes suppress DSB formation. Fig. 3C shows that both full-length hNTH1 and the high turnover truncation mutant hNTH1Δ55 produced DSBs with similar efficiencies. Thus, the relatively poor lyase activity of hNTH1, but not its slow turnover rate, was limiting for DSB production in cases when opposing lesions were separated by 5 bp. In cells, hAPE1 is thought to be present in excess over individual DNA glycosylases (Chen et al., 1991; Kane & Linn, 1981). We therefore included hAPE1 in most subsequent experiments, reasoning that doing so would better approximate in vivo conditions.

**DSB formation due to processing of clustered lesions in 601 nucleosomes.**

The oxidative lesions in the 5S rDNA-based nucleosomes used for the above-described experiments were located relatively close (~12 to ~37 nt depending on translational position) to the nucleosome edge. To determine if DNA glycosylases
generate DSBs when processing clustered lesions located closer to the dyad axis, and to control for possible sequence context effects, we replaced two thymine residues in 601 DNA with Tg, such that, when assembled into a nucleosome, the top and bottom strand lesions would lie 39 and 44 nt from the dyad axis (Chua et al., 2012). We predicted that both lesions in the resulting “Tg39+5” nucleosome would be only marginally accessible to hNTH1, and thus expected to see very limited DSB production. Surprisingly, the efficiency of DSB production in the 601 nucleosomes was as high as that observed for Tg51-5 nucleosomes (Figs. 4A & B).

Fig. 4C documents the effect of increasing glycosylase concentration on the efficiency of DSB formation. For enzyme concentrations in excess over total lesion (1 nM), the efficiency of DSB formation in naked DNA quickly reached a plateau value. In contrast, the efficiency of DSB formation in nucleosomes continued to climb over the enzyme range tested. This increasing yield of DSBs at high enzyme concentrations is a hallmark associated with the capture of lesions that are accessible only during periodic unwrapping events (Maher et al., 2013; Prasad, Wallace, et al., 2007). The fact that this phenomenon was more pronounced for the 601 Tg39+5 nucleosomes than for the 5S rDNA nucleosomes is consistent with the greater distance between the lesions in Tg39+5 and the nucleosome edge, and a correspondingly lower rate of unwrapping-mediated lesion exposure.

**Evidence for coupled processing of clustered lesions in nucleosomes.**
Given the increased distance between the lesions in 601 Tg39+5 nucleosomes and the nucleosome edge, the production of DSBs was surprisingly efficient. This suggested that the two lesion processing events required for DSB formation might somehow be coupled. If instead opposing strand lesions are processed independently of one another, the frequency of DSB formation (measured using non-denaturing gels such as that shown in Fig. 5A) should equal the product of the two single strand cleavage frequencies (measured using denaturing gels such as that shown in Fig. 5B). To distinguish between independent and coupled processing of lesions, we measured rates of single-strand break formation for both the lesion-containing 5S rDNA nucleosomes and naked DNA controls (Figs. 5C & D and Table 1), and for the above-described 601 nucleosomes (Fig. 5E and Table 1). In the case of naked, Tg51-5 DNA substrates, Fig. 5C shows that hNTH1 processed ~80% of the lesions in each DNA strand, leading to formation of DSBs in (0.8*0.8) ~64% of molecules within the reaction timeframe. Table 2 compares observed DSB frequencies from multiple experiments to those predicted from single-strand cleavage data, assuming independent processing of single lesions. If this assumption is correct, a plot of the observed versus predicted values will produce a slope equal to 1.0. For naked DNA substrates, Table 2 reports slopes that range from 0.8 to 1.0. This is consistent with the independent processing assumption or a slightly higher efficiency of lesion processing in fully intact DNA.

Although hNTH1 processed both lesions in naked Tg51-5 DNA, it exhibited a ~3-fold preference for processing the top strand lesion in nucleosomes (Fig. 5D and Table 1). Thus, the helical orientation of the top strand lesion is closer to optimal than that of the
bottom strand lesion. If the processing of the top and bottom lesions occurred independently of each other, the yield of DSBs would be \((0.85 \times 0.33) \approx 28\%\). This prediction falls within the margin of error for the observed DSB frequency of 33±11\%, which is consistent with an independent processing model.

In contrast to the results obtained for the 5S rDNA-based nucleosomes, we observed significant differences between the expected and observed rates of DSB formation in the 601-based Tg39+5 nucleosomes. Specifically, both top and bottom strand lesions in Tg39+5 nucleosomes were processed with similar efficiencies (Fig. 5E and Table 1), and the efficiency of DSB formation was more than double that expected for independent processing events (Table 2). The simplest explanation for this observation is that one of the two lesions in Tg39+5 is bound by hNTH1 during a transient DNA unwrapping event, and that this binding stabilizes the unwrapped configuration long enough to facilitate the binding of a second DNA glycosylase to the opposing strand lesion.

**The efficiency of BER-dependent formation of DSBs in nucleosomes varies among different DNA glycosylases.**

To determine if DNA glycosylases other than hNTH1 generate DSBs during the repair of clustered lesions, we conducted similar experiments with the human 8-oxoguanine glycosylase (hOGG1). As with hNTH1, hOGG1 is a bifunctional enzyme, possessing both a glycosylase (which removes 7,8-dihydro-8-oxoguanine (8-oxoG) lesions (Leipold et al., 2003)) and a DNA lyase activity. As before, we designed
constructs with the goal of disproving the hypothesis that nucleosomes fully suppress DSB formation. To minimize the possibility of altering nucleosome positioning determinants, we replaced two, naturally occurring guanosine residues in 5S rDNA with 8-oxoG, forming the construct 8oxoG50-5. Compared to Tg51-5, the 8oxoG50-5 lesions were positioned one bp closer to the dyad axis, and differed in helical orientation by ~36° relative to the underlying histone octamer. Surprisingly, hOGG1 produced very few DSBs in naked 8oxoG50-5 DNA and virtually none in nucleosomes, even when used at very high concentrations (e.g. 500 nM active enzyme, a 1000-fold excess over substrate) (Figs. 6A & B). As with hNTH1, hAPE1 had been reported to stimulate the glycosylase activity of hOGG1 (Hill et al., 2001; Vidal et al., 2001). Although hAPE1 is active only in the presence of magnesium, it will displace hOGG1 from hOGG1-generated abasic sites, even in the absence of magnesium. Thus, if binding of hOGG1 to one lesion were interfering with binding to the opposing strand lesion, the addition of hAPE1 with no magnesium would have enhanced DSB formation. It did not. Instead, both magnesium and hAPE1 were required for efficient hOGG1-initiated DSB formation in naked DNA (Fig. 6B, top graph).

While the addition of magnesium and hAPE1 substantially increased the hOGG1-initiated production of DSBs in naked DNA, it only marginally increased production of DSBs in nucleosomes (Fig. 6B, bottom graph). As a result, hOGG1- and hAPE1-dependent production of DSBs in naked DNA was nearly as high as that observed in reactions with hNTH1 and hAPE1 (Table 2), while hOGG1- and hAPE1-dependent production of DSBs in nucleosomes remained far lower than that produced by hNTH1.
(7±1% vs. 33±7% DSBs, respectively). To determine if this result could be attributed to the nucleosome substrate, we treated the 8oxoG50-5 nucleosomes with the *E. coli* glycosylase Fpg. Fpg is a functional homolog to hOGG1 and, importantly, did not exhibit the strand preference in naked DNA that we had observed for reactions containing hOGG1 and hAPE1 (Table 1). Fig. 6C shows that 50 nM Fpg alone generated about three-fold more DSBs in naked DNA than did hOGG1 together with hAPE1. As with hOGG1 and hNTH1, Fpg exhibited a clear strand preference when processing lesions in nucleosomes (Fig. 6D); this preference could be attributed to the differences in helical orientation of the opposing strand lesions.

The reasons for the very low efficiency of hOGG1-mediated DSB formation in nucleosomes became apparent after careful quantification of single lesion processing efficiencies. In nucleosomes, hOGG1 is subject to steric constraints that are similar to those evident for Fpg. However, hOGG1 sequence preferences severely limited processing of lesions in the more accessible of the opposing strand lesions in 8oxoG50-5 nucleosomes (Table 1). Meanwhile, the preferred bottom strand lesion was sterically occluded (Fig. 6D). Thus, an unfavorable helical orientation in nucleosomes trumped whatever attributes account for preferential processing of the bottom strand lesion in naked DNA.

**Fate of nucleosomes following DSB formation.**

Most DSBs in cells are believed to be quickly bound by Ku proteins and channeled into DSB repair pathways. However, the detection and processing of a DSB in
nucleosomal DNA might be influenced by whether the nucleosome remains intact after DSB formation. Therefore, to assess the fate of nucleosomes following DSB formation, we treated 5S Tg51-5 and 601 Tg39+5 nucleosomes with hNTH1 and hAPE1 in separate reactions, for 30 minutes at 37°C, and then immediately loaded aliquots from the reactions onto native gels (Fig. 7). Lane 4 in Fig. 7A shows the shorter of the two DNA fragments created by DSB formation in 5S Tg51-5 nucleosomes migrating independently of the nucleosome. Thus, one end of the hNTH1-generated DSB was released, while the other end remained nucleosome-associated. By contrast, little if any DNA was released following DSB formation in 601 Tg39+5 nucleosomes (Fig. 7A, lane 9). In both cases, however, DSB formation altered the mobility of the residual nucleosome population. To determine if these motility changes were due to altered histone-DNA interactions, we incubated Tg51-5 nucleosomes with hNTH1 and hAPE1 as before but in the presence of a large excess of Age I. The Age I restriction site lies well-within the 5S rDNA nucleosome (Fig. 7B) and we had previously determined that this site is almost completely occluded in nucleosomes (Prasad, Wallace, et al., 2007). As shown in Fig. 7C, hNTH1 and hAPE1 generated DSBs in ~30% of all nucleosomes. By contrast, Age I cleaved 9.4% of the nucleosomal sites in the absence of hNTH1 and hAPE1, and only 11.6% of the nucleosomal sites in the presence of hNTH1 and hAPE1. These results indicate that formation of DSBs at the sites tested does not invariably result in nucleosome disruption. Whether nucleosomes sequester DSB ends once they form probably depends on the distance between the cleavage site and the dyad axis.
2.5 Discussion

Exposure to ionizing radiation, either from X-rays, radioactive elements, or naturally occurring cosmic radiation, can result in potentially lethal DSBs in DNA. DSBs can occur through direct interactions between photons and the DNA backbone, or indirectly, through formation of clustered, oxidative lesions, followed by attempted BER. The studies in this paper indicate that the packaging of DNA into nucleosomes substantially suppresses BER-dependent formation of DSBs. The degree of suppression depends on multiple variables, including the location of opposing strand lesions relative to one another, their helical orientation relative to the underlying histone octamer, and differences in the glycosylases that initiate BER.

To what extent do nucleosomes protect cells from BER-dependent DSB formation?

As noted in the Introduction, lesion recognition and early steps in short patch BER do not invariably require or lead to the movement or disruption of nucleosomes. Hence, the nucleosome-dependent suppression of DSB formation documented in this study may occur in vivo as well. In addition to the many factors that influence rates of DSB formation in model nucleosomes, there are even more that probably influence DSB formation in cells. Hence, any numerical estimate of the magnitude of protection from DSB formation that nucleosomes offer must be somewhat speculative. With that caveat, enumerating nucleosome-related protective factors may be of heuristic value. First, nucleosomes offer limited protection from lesion formation. Hydroxyl radical footprinting of nucleosomal DNA typically reveals a ~10 bp pattern of susceptibility to
hydroxyl radical-mediated damage (Hayes et al., 1991). This pattern reflects both the periodic expansion and compression of major and minor grooves in DNA as it is wound about the histone octamer, and discrete interactions every ~10 bp between DNA and the histone octamer (Luger et al., 1997). However, the difference between highly and poorly reactive residues in nucleosomes is modest, on the order of two-three fold. Therefore, on average, nucleosomes provide no more than a 2-fold protection from hydroxyl-radical mediated damage compared to naked DNA. The probability that two oxidative lesions will form close to one another in opposing strands in nucleosomes would thus be about \((0.5)^2\) relative to that in naked DNA. Second, this study has reinforced earlier observations that BER-dependent formation of DSB requires lesions to be offset from one another by at least three bp. In naked DNA, the maximum offset will likely be governed by the size of oxidative lesion clusters, which in turn will reflect the reactive lifetime and diffusion rates of hydroxyl radicals. In nucleosomes the maximum offset is likely governed by steric factors that influence binding of DNA glycosylases. For example, the helical orientation chosen for one of the lesions in Tg51-5 was close to optimal but much less so for the opposing strand lesion, with the result that DSB formation was suppressed 3-4 fold compared to naked DNA. Thus, tighter constraints on lesion stagger will reduce the number of lesion configurations that could lead to DSB formation in nucleosomes as compared to naked DNA.

The above-described factors combined suggest nucleosomes suppress BER-mediated DSB formation by at least 10-fold. The contributions of other factors are harder to quantify. For example, lesions at occluded sites in nucleosomes are made transiently
accessible by periodic unwrapping of DNA from the edge of the histone octamer. The unwrapping mediated exposure of clustered lesions appears to occur at a biologically meaningful rate for lesions located within 30-40 bp of either edge of the nucleosome but not for more centrally located lesions (Maher et al., 2013). This phenomenon, along with the coupled lesion processing that we observed in the Tg39+5 nucleosomes, is likely to reduce the overall protection that nucleosomes offer against DSB formation. In cells, such unwrapping events, in some cases assisted or amplified by chromatin remodeling agents, may be confined to regions devoid of linker histones (Horn et al., 2002). A second factor that may also increase the vulnerability of DNA near the nucleosome edges to DSB formation relates to the fact that many nucleosomes occupy multiple overlapping translational positions. As noted earlier, this is the case for the 5S rDNA nucleosomes used in the present study, and also appears to be the case for nucleosomes in budding yeast (Cole et al., 2012). However, in more complex eukaryotes histone H1 may reduce positional mobility (Hill & Imbalzano, 2000).

In summary, chromatin associated with linker histones, heterochromatic regions associated with multiple repressive factors, and highly condensed chromatin in cells undergoing mitosis or meiosis probably limit BER-dependent DSB formation to an even greater extent than documented in this study. Indeed, chromatin condensation has been reported to protect DNA from damaging radiation, and decondensation (not necessarily transcription-related) is sufficient to sensitize DNA to radiation (Falk et al., 2008; Takata et al., 2013). A proposed mechanism for this phenomenon is that condensation of chromatin helps to exclude water and thus fewer ROSs are generated near DNA as a
result of irradiation. An alternative or additional mechanism is that BER machinery is unable to detect or act on lesions in highly condensed chromatin, thereby reducing the abundance of DSBs formed. On the other hand, chromatin decondensation and nucleosome remodeling during replication and transcription may increase not only the frequency with which clustered lesions form but also the efficiency of BER and the chance that concurrent BER reactions will generate DSBs. This prediction could explain why continuously transcribed regions of genomic DNA are unusually susceptible to damage induced by ionizing radiation [reviewed in (Bailey & Bedford, 2006)].

2.6 References


from chromatin by the human DNA glycosylase, NEIL1. *DNA Repair (Amst)*, 9(2), 134-143.


### 2.7 Footnotes

**Acknowledgements.**

We thank April Averill for enzyme purifications, Andrea Lee for the plasmid encoding hOGG1, and Robyn Maher for critical comments on the manuscript.

**Funding.**

This research was supported by NSF grant MCB-0821941 and NCI grant P01-CA098993 to D.S. Pederson. The authors of this study have no financial conflicts of interest.
Abbreviations used.

The abbreviations used are: 8-oxoguanine or 8-oxoG, 7,8-dihydro-8-oxoguanine; AP site, apurinic/apyrimidinic site; BER, base excision repair; bp, base pair; DSB, double-strand break; Fpg, *E. coli* formamidopyrimidine-DNA glycosylase; hAPE1, human apurinic/apyrimidinic endonuclease 1; hNTH1, human endonuclease III; hOGG1, human 8-oxoguanine glycosylase; oligo, oligodeoxyribonucleotide; nt, nucleotide; ROS, reactive oxygen species; SSB, single-strand break; Tg, thymine glycol.

2.8 Table and Figure Legends

**TABLE 1. The nucleosome suppresses single-strand break formation relative to naked DNA.** The abundance of single strand DNA breaks generated during 30 min reactions with the indicated enzymes and templates is tabulated. Substrate concentrations in all experiments were 0.5 nM. As noted in Experimental Procedures, reaction endpoints were either determined in triplicate or by fitting multiple time points to single exponential curves. For clarity, standard deviations are omitted. Nucleosome-dependent suppression of strand cleavages was calculated by dividing naked DNA cleavage frequency by its corresponding nucleosome strand cleavage frequency, using data from reactions where single-strand cleavage extents were <90% (i.e. where cleavage frequencies were less than maximal).
**TABLE 2. Predicted vs. observed double-strand break formation.** The predicted yield of DSBs, assuming that processing of lesions on opposing strands occurred independently of one another, was calculated by multiplying the SSB frequency of bottom and top strands, using data from Table 1. The predicted and observed yield of DSBs were then plotted against one another, and fit to a linear regression with both x- and y-intercepts set to zero. Slope and $R^2$ values are shown. The slopes for all of the naked DNA substrates shown and many of the nucleosome substrates were approximately equal to 1.0, indicating a close correspondence between the predicted and observed DSB frequencies. A slope much less than 1.0 (i.e. observed DSB frequencies are lower than predicted) would have indicated that cleavage at one strand hinders cleavage of the opposite strand. A slope much greater than 1.0 (i.e. observed DSB frequencies are higher than predicted) would suggest that cleavage at one strand enhances cleavage of the opposite strand. This was the case for hNTH1 acting on lesions in 601 nucleosomes.

**FIGURE 1. Double-strand break formation during attempted base excision repair (BER).** (A) Schematic of attempted BER resulting in DSB formation. Ionizing radiation causes multiple, closely spaced oxidative lesions on opposing DNA strands. Near-simultaneous strand cleavages by either a glycosylase or AP-endonuclease result in a DSB that can no longer be repaired by the BER pathway. (B) Summary of 5S rDNA-based nucleosome constructs, indicating the approximate position of lesions with respect to the histone core (H2A-H2B and H4-H3 dimers). Construct names are in parentheses.
and indicate: the lesion (Tg = thymine glycol; 8-oxoG = 8-oxoguanine), top strand lesion position with respect to the dyad axis (-50 or -51), and base pair stagger (2 or 5) between opposing strands. The predicted yield of DSBs is based upon previous work using naked DNA substrates (Harrison et al., 1998; Harrison et al., 1999) as well as relative accessibility of the lesions with respect to the histone core. (C) Crystal structure of the 1.9Å alpha-satellite nucleosome (PDB 1KX5), oriented to highlight the 51-2 and 51-5 lesion positions. Position -51 is located on the orange strand, with 2 bp and 5 bp opposing lesions on the purple strand, closer towards the dyad axis (lesion positions are in blue and circled). Lesion orientations in 8oxoG50-5 differ from those shown for Tg51-5 by ~36°, whereas those for Tg39+5 differ by ~72°. The glycosylase must bind to the lesion via the minor groove, and then "flip" the damaged base into its active site through the major groove.

**FIGURE 2. Impact of the stagger between opposing lesions on double-strand break formation.** (A) Representative nucleosome preparations examined by native PAGE. (B) Representative native gel to assess DSB formation following incubation of 0.5 nM Tg51-2 DNA and nucleosomes with 25 nM hNTH1Δ55 at 30°C. Aliquots were removed at the indicated time points, quenched, and analyzed using 6% native PAGE (see Experimental Procedures). A DSB will result in left (L) and right (R) product bands (boxed). Poor DSB formation is observed in a 2 bp staggered construct. (C) Representative DSB gel showing hNTH1-catalyzed DSB formation in Tg51-5 DNA and nucleosomes treated with hNTH1Δ55 as in Fig. 2B. DSB product bands are boxed. (D) Quantification of DSB
formation comparing the 2 bp stagger to the 5 bp stagger. Error bars represent standard deviations (n=3). Nucleosomes exhibit pronounced protection from DSBs compared to naked DNA.

FIGURE 3. hNTH1-mediated, double-strand break formation in nucleosomes is lyase-limited. (A) Representative gel showing DSB formation in 0.5 nM Tg51-5 nucleosomes and naked DNA by 25 nM hNTH1 +/- 25 nM human apurinic/apyrimidinic endonuclease 1 (hAPE1). As in Fig. 2B, DSB product bands (L and R) are boxed. (B) Quantification of DSB formation in naked DNA (top) and nucleosomes (bottom) in the presence and absence of hAPE1. Error bars represent standard deviation (n=4). hAPE1 addition results in a visible increase in DSB formation rate. (C) Wild type recombinant hNTH1 compared with a truncation mutant with a higher catalytic turnover rate (hNTH1Δ55). See Figs. 3A and 2C, respectively, for representative DSB gels of each enzyme. Graphs represent DSB formation in the Tg51-5 construct: free DNA (left) and nucleosomes (right). The truncation mutant behaves similarly to the full length protein. Error bars represent standard deviations (n=3).

FIGURE 4. DSB formation due to processing of clustered lesions in 601 nucleosomes. (A) Representative nucleosome preparation (left) and DSB gel assay (right) for Tg39+5, which contains two thymine glycol (Tg) lesions on opposing strands, spaced 5 bp apart. DSB assays were performed with 0.5 nM lesion-containing substrate and 25 nM hNTH1 as per Experimental Procedures. Product bands are boxed as before.
(B) Despite the multiple differences between Tg39+5 and Tg51-5, hNTH1 produced similar numbers of DSBs in both naked DNA and nucleosome substrates (5S Tg51-5 data from Fig. 3B are repeated here for comparison). (C) Effect of varying enzyme concentrations on the yield of DSBs in Tg39+5 DNA and nucleosomes as compared to the DSB yield in 5S Tg51-5 substrates under identical reaction conditions. Assays were performed with 0.5 nM substrate and 0.5, 2.5, or 10 nM hNTH1 + 25 nM hAPE1. Naked DNA DSB formation (left) plateaus quickly. 601 Tg39+5 nucleosomal DNA (right), however, shows an increase in DSBs with enzyme concentration, indicating increased binding, which in turn suggests that transient nucleosome unwrapping creates a limiting effect on DSB formation. This effect is not as pronounced in the 5S Tg51-5 nucleosomal DNA, which may be due in part to the 5S lesions being located farther away from the dyad axis and/or multiple, minor translational positions placing some lesion populations even farther from the dyad. Error bars represent standard deviations (n=3).

FIGURE 5. Evidence for coupled processing of clustered lesions in nucleosomes. 5S Tg51-5 and 601 Tg 39+5 constructs (0.5 nM) were treated with varying hNTH1 enzyme concentrations, as per the Experimental Procedures. (A) Representative hNTH1-catalyzed gel showing DSB formation in Tg51-5 substrates. (B) Representative denaturing gel showing SSBs generated in the same reaction shown in Fig. 5A. The DNA band marked with an asterisk represents an incompletely ligated substrate; it does not contain a lesion and thus had no impact on results. (C) Quantification of single- and double-strand break formation in 5S Tg51-5 naked DNA. In naked DNA, DSB formation equals the products
the product of the single-strand break frequencies which are equal for both DNA strands, whereas in 5S Tg51-5 nucleosomes (D) the rate of DSB formation is limited by the production of SSBs in the more occluded of the two DNA strands. (E) Unlike Tg51-5 results, 601 Tg39+5 nucleosome strand cleavage was approximately equal for each strand, suggesting that neither strand specifically limits DSB formation. Furthermore, the observed rate of DSB formation is greater than predicted, based on SSB results (Table 2). This suggests that processing of one lesion may facilitate binding and processing of the second lesion. Error bars represent standard deviations (n=3).

FIGURE 6. The efficiency of BER-dependent formation of DSBs in nucleosomes varies substantially among different DNA glycosylases. (A) Representative nucleosome preparation (top) and DSB gel (bottom) for 8oxoG50-5, containing two 8-oxoguanine (8-oxoG) lesions on opposing strands, spaced 5 bp apart. 8oxoG50-5 DNA or nucleosomes (0.5 nM) were incubated with 500 nM hOGG1 + 25 nM hAPE1 as per Experimental Procedures. DSB products are boxed. (B) DSB formation with hOGG1 was assessed under multiple conditions (+/- 25 nM hAPE1, +/- 5 mM MgCl2) in both naked DNA (top) and nucleosomes (bottom). hAPE1 and Mg2+ are both required for DSB formation in naked DNA, indicating that the lyase activity of hOGG1 is limiting for DSB formation. Even at high hOGG1 concentrations with hAPE1 and Mg2+, DSB formation in nucleosomes is severely limited for reasons discussed in the text. (C) DSB formation in 8oxoG50-5 substrates by 50 nM Fpg compared to 50 nM hOGG1. In contrast to hOGG1, Fpg readily forms DSBs in both naked and nucleosomal DNA. (D) and (E) SSB data for
FIGURE 7. Fate of nucleosomes during the processing of closely opposed oxidative lesions. (A) Native gel shifts showing 0.5 nM naked and nucleosome DNA before (Lanes 1, 2, 6, & 7) and after (Lanes 3, 4, 8 & 9) treatment of 5S Tg51-5 substrates (left) and Tg39+5 substrates (right) with 25 nM hNTH1 and 25 nM hAPE1. Samples were run on 6% native PAGE without protein removal. Formation of DSBs on the nucleosome can result in release (dotted box, lane 4) or retention (lane 9) of cleaved DNA, depending upon sequence and/or proximity of the DSB to the dyad axis. (B) Restriction digest map of the 5S DNA substrate, indicating fragment sizes from either DSB formation or Age I cleavage. The large, grey oval represents the portion of DNA that is associated with the nucleosome. (C) Age I nucleosome disruption assay. Naked DNA or nucleosomes were assayed with 25 nM hNTH1, 25 nM hAPE1, and 50 U AgeI (Lanes 4 & 8) as per Experimental Procedures. DSB formation can be observed by R and L bands. Nucleosome disruption was assessed by calculating the increase in Age I cleavage (71 bp band “B” in Lanes 7 & 8) before and after DSB formation. The Age I site remains largely inaccessible following hNTH1/hAPE1 addition, indicating an intact, undisturbed nucleosome. To quantify results, we tracked the 71 bp Age I cleavage product (see Fig. 7B for a list of all products) and calculated cleavage efficiency as product (doubled) divided by the sum of all bands. Nucleosome Age I cleavage was then expressed as a fraction of total naked DNA cleavage.
2.9 Tables and Figures

Table 2-1. The nucleosome suppresses single-strand break formation relative to naked DNA.

<table>
<thead>
<tr>
<th>Glycosylase &amp; substrate</th>
<th>Glycosylase nM</th>
<th>APE? (n)</th>
<th>DNA / nucleosomes</th>
<th>Fold suppression*</th>
<th>DNA / nucleosomes</th>
<th>Fold suppression*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>500</td>
<td>N 1</td>
<td>0.42 / 0.04</td>
<td>11.2</td>
<td>0.26 / 0.14</td>
<td>1.8</td>
</tr>
<tr>
<td>hOGG1 + 8oxoG50-5</td>
<td>50</td>
<td>Y 1</td>
<td>0.87 / 0.17</td>
<td>5.1</td>
<td>0.38 / 0.24</td>
<td>1.6</td>
</tr>
<tr>
<td></td>
<td>200</td>
<td>Y 1</td>
<td>0.99 / 0.24</td>
<td>-</td>
<td>0.54 / 0.39</td>
<td>1.4</td>
</tr>
<tr>
<td></td>
<td>500</td>
<td>Y 2</td>
<td>0.92 / 0.22</td>
<td>-</td>
<td>0.69 / 0.42</td>
<td>1.6</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>N 3</td>
<td>0.76 / 0.23</td>
<td>3.3</td>
<td>0.65 / 0.72</td>
<td>0.9</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>N 3</td>
<td>0.81 / 0.28</td>
<td>2.9</td>
<td>0.78 / 0.85</td>
<td>0.9</td>
</tr>
<tr>
<td>hNTH1 + Tg51-5</td>
<td>125</td>
<td>N 3</td>
<td>0.81 / 0.33</td>
<td>2.5</td>
<td>0.84 / 0.85</td>
<td>1.0</td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td>Y 3</td>
<td>0.87 / 0.24</td>
<td>3.7</td>
<td>0.64 / 0.53</td>
<td>1.2</td>
</tr>
<tr>
<td></td>
<td>2.5</td>
<td>Y 3</td>
<td>0.88 / 0.33</td>
<td>2.6</td>
<td>0.66 / 0.60</td>
<td>1.1</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>Y 3</td>
<td>0.93 / 0.41</td>
<td>-</td>
<td>0.71 / 0.64</td>
<td>1.1</td>
</tr>
<tr>
<td></td>
<td>25</td>
<td>N 1</td>
<td>0.87 / 0.21</td>
<td>4.1</td>
<td>0.80 / 0.36</td>
<td>2.2</td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td>Y 3</td>
<td>0.83 / 0.15</td>
<td>5.6</td>
<td>0.69 / 0.15</td>
<td>4.6</td>
</tr>
<tr>
<td>hNTH1 + Tg39+5</td>
<td>2.5</td>
<td>Y 3</td>
<td>0.89 / 0.23</td>
<td>3.9</td>
<td>0.75 / 0.24</td>
<td>3.1</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>Y 3</td>
<td>0.80 / 0.36</td>
<td>-</td>
<td>0.70 / 0.35</td>
<td>-</td>
</tr>
</tbody>
</table>

*Fold-suppression data calculated only where ss cleavage frequencies for both nucleosome and naked DNA controls were < 90%.
Table 2-2. Predicted vs. observed double-strand break formation.

<table>
<thead>
<tr>
<th>Glycosylase &amp; substrate</th>
<th>nM</th>
<th>25 nM APE? (n)</th>
<th>Predicted DSB frequency**</th>
<th>Observed DSB frequency</th>
<th>Slope ($R^2$) for observed vs. predicted DSB frequencies</th>
<th>Predicted DSB frequency**</th>
<th>Observed DSB frequency</th>
<th>Slope ($R^2$) for observed vs. predicted DSB frequencies</th>
</tr>
</thead>
<tbody>
<tr>
<td>hOGG1 + 8oxoG50-5</td>
<td>500</td>
<td>N 1</td>
<td>0.11</td>
<td>0.09</td>
<td>0.9 (0.99)</td>
<td>0.01</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>Y 1</td>
<td>0.33</td>
<td>0.29</td>
<td></td>
<td>0.04</td>
<td>0.06</td>
<td>1.1 (0.85)</td>
</tr>
<tr>
<td></td>
<td>200</td>
<td>Y 1</td>
<td>0.53</td>
<td>0.47</td>
<td></td>
<td>0.09</td>
<td>0.13</td>
<td></td>
</tr>
<tr>
<td></td>
<td>500</td>
<td>Y 2</td>
<td>0.63</td>
<td>0.52</td>
<td></td>
<td>0.09</td>
<td>0.07</td>
<td></td>
</tr>
<tr>
<td>hNTH1 + Tg51-5</td>
<td>10</td>
<td>N 3</td>
<td>0.49</td>
<td>0.44</td>
<td></td>
<td>0.17</td>
<td>0.22</td>
<td></td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>N 3</td>
<td>0.64</td>
<td>0.55</td>
<td></td>
<td>0.24</td>
<td>0.27</td>
<td></td>
</tr>
<tr>
<td></td>
<td>125</td>
<td>N 3</td>
<td>0.68</td>
<td>0.57</td>
<td>0.8 (0.96)</td>
<td>0.28</td>
<td>0.33</td>
<td>1.2 (0.77)</td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td>Y 3</td>
<td>0.56</td>
<td>0.46</td>
<td></td>
<td>0.13</td>
<td>0.18</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2.5</td>
<td>Y 3</td>
<td>0.58</td>
<td>0.50</td>
<td></td>
<td>0.20</td>
<td>0.23</td>
<td></td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>Y 3</td>
<td>0.66</td>
<td>0.56</td>
<td></td>
<td>0.26</td>
<td>0.27</td>
<td></td>
</tr>
<tr>
<td>hNTH1 + Tg29+5</td>
<td>25</td>
<td>N 1</td>
<td>0.69</td>
<td>0.67</td>
<td></td>
<td>0.08</td>
<td>0.19</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td>Y 3</td>
<td>0.58</td>
<td>0.58</td>
<td>1.0 (0.74)</td>
<td>0.02</td>
<td>0.08</td>
<td>2.1 (0.72)</td>
</tr>
<tr>
<td></td>
<td>2.5</td>
<td>Y 3</td>
<td>0.67</td>
<td>0.62</td>
<td></td>
<td>0.05</td>
<td>0.14</td>
<td></td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>Y 3</td>
<td>0.56</td>
<td>0.62</td>
<td></td>
<td>0.12</td>
<td>0.23</td>
<td></td>
</tr>
</tbody>
</table>

*Equal to the product of the top and bottom strand SSB frequencies: assumes SSB formation events are independent.

**Data in italics omitted from linear regression calculations.
Figure 2-1. Double-strand break formation during attempted base excision repair (BER).

### A

- **Damaged base**
- **Ionizing radiation** → **Glycosylase-mediated excision of damaged base** → **Glycosylase or AP-mediated lyase** → **Double-strand break (DSB)**

### B

<table>
<thead>
<tr>
<th>Stagger between opposing lesions</th>
<th>Predicted DSB Yield</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 bp (&quot;Tg51-2&quot;)</td>
<td>Naked DNA</td>
</tr>
<tr>
<td></td>
<td>DNA</td>
</tr>
<tr>
<td>H2A-H2B</td>
<td>Poor</td>
</tr>
<tr>
<td>H4-H3</td>
<td>Poor</td>
</tr>
<tr>
<td>5 bp (&quot;Tg51-5&quot;)</td>
<td>Nucleosomes</td>
</tr>
<tr>
<td></td>
<td>Good</td>
</tr>
<tr>
<td>H2A-H2B</td>
<td>Fair</td>
</tr>
<tr>
<td>H4-H3</td>
<td></td>
</tr>
</tbody>
</table>

= oxidative base damage

### C

- **Major Groove**
- **Minor Groove**
- **-46**
- **-49**
- **51**

Figure 2-1. Double-strand break formation during attempted base excision repair (BER).
Figure 2-2. Impact of the stagger between opposing lesions on double-strand break formation.
Figure 2-3. hNTH1-mediated, double-strand break formation in nucleosomes is lyase-limited.
Figure 2-4. DSB formation due to processing of clustered lesions in 601 nucleosomes.
Figure 2-5. Evidence for coupled processing of clustered lesions in nucleosomes.
Figure 2-6. The efficiency of BER-dependent formation of DSBs in nucleosomes varies substantially among different DNA glycosylases.
Figure 2-7. Fate of nucleosomes during the processing of closely opposed oxidative lesions.
CHAPTER 3: HUMAN DNA LIGASE IIIα FORMS A REVERSIBLE, LIGATION-PERMISSIVE COMPLEX WITH DNA NICK-CONTAINING NUCLEOSOMES

Wendy J. Cannan¹, Alan E. Tomkinson², Susan S. Wallace¹ and David S. Pederson¹

¹Department of Microbiology & Molecular Genetics, University of Vermont, Stafford Building, Room 220, 95 Carrigan Drive, Burlington, VT, 05405, USA, Tel.: (802) 656-8587; E-mail: david.pederson@uvm.edu

²The University of New Mexico Cancer Center, and the Departments of Internal Medicine & Molecular Genetics & Microbiology, University of New Mexico, Albuquerque, NM 87131, USA

3.1 Abstract

Reactive oxygen species generate potentially cytotoxic and mutagenic lesions in DNA, both between and within the nucleosomes that package DNA in chromatin. Enzymes that catalyze early steps in the base excision repair (BER) of oxidative lesions can process lesions at many sites in nucleosomes without the aid of chromatin-remodeling agents, and without irreversibly disrupting the host nucleosome. However, it was not known if host nucleosomes survive the final step in BER, namely the sealing of DNA nicks by DNA ligase IIIα-XRCC1. We present evidence that periodic, spontaneous
partial unwrapping of DNA from the histone octamer facilitates the nick discovery step, and that upon binding to a DNA gap or nick ligase IIIα-XRCC1 takes on a relatively salt-resistant configuration. Ligation occurs within a ternary complex that ligase IIIα-XRCC1 forms with nick-containing nucleosomes. Once ligation is complete, the ternary complex decays and, remarkably, host nucleosomes spontaneously return to their native configurations. We also find that the efficiency of ligation varies inversely with nucleosome stability, which suggests that ligase IIIα-XRCC1 may drive nucleosomes into a ligation-permissive configuration, through folding into an active configuration.

3.2 Introduction

Every day, reactive oxygen species, produced during normal oxidative metabolism and exposure to ionizing radiation, generate some ~30,000 oxidized bases and abasic sites in the DNA of every nucleated human cell (Fraga et al., 1990; Lindahl & Nyberg, 1972); reviewed in (Friedberg et al., 2005)). Left unrepaired, these oxidative damages can lead to polymerase stalling, DNA strand breaks, base mutations, and cell death. The vast majority of these oxidative lesions are subject to base excision repair (BER; reviewed in (David et al., 2007; Duclos et al., 2012; Hegde, Hazra, et al., 2008; Krokan & Bjoras, 2013; Robertson et al., 2009; Wallace, 2014)), which begins with the excision of an oxidized base by a DNA glycosylase, followed by cleavage of the DNA backbone at the resulting abasic site, either by a glycosylase-associated lyase activity or by apurinic endonuclease (APE1). This leaves a polymerase-blocking moiety on the 3’ side, which is removed by either APE1 or polynucleotide kinase phosphatase. The result
is a single strand deoxynucleotide gap, flanked by a 3' hydroxyl group and, in some cases, a 5'-2-deoxyribose phosphate (dRP) moiety. In short-patch BER, DNA polymerase β (pol β) fills the DNA gap and removes the blocking dRP group, leaving a nick that is sealed by DNA ligase IIIα, in association with the scaffolding protein XRCC1 (X-ray repair cross-complementing protein 1). If pol β fails to remove the 5' moiety, long patch BER ensues. This entails recruitment of a replicative polymerase that displaces DNA linked to the 5' moiety. The resulting DNA flap is removed by flap endonuclease 1, leaving a DNA nick that is sealed by DNA ligase I.

In eukaryotes, BER occurs in chromatin, where most of the DNA is wrapped around histone octamers to form nucleosomes (Luger et al., 1997). The close association between DNA and histone octamers, and the changes in major and minor groove geometry associated with DNA bending about the octamer, impedes the binding of many regulatory factors and enzymes. These impediments are sometimes relieved by chromatin remodeling agents (reviewed in (Clapier & Cairns, 2009)). The nucleosome remodeling complexes SWI/SNF, ISW1, and ISW2 enhance the activity of certain BER enzymes on nucleosome substrates in vitro (Menoni et al., 2007; Nakanishi et al., 2007) and, in budding yeast, depletion of an essential subunit in the chromatin remodeling complex RSC produces numerous DNA and chromatin-related defects, among them defects in BER (Czaja et al., 2014). However, it is not yet clear that chromatin-remodeling agents act at sites of oxidative damage in cells. If chromatin-remodelers are found to promote BER in vivo, it will be important to determine when and how they are recruited to sites of oxidative damage. To identify chromatin-associated, rate limiting steps in BER, and
possible “pioneer factors” that may recruit chromatin-remodeling agents, we have investigated interactions between BER factors and nucleosomes (Cannan et al., 2014; Maher et al., 2013; Odell et al., 2011; Odell et al., 2010; Prasad, Wallace, et al., 2007). The DNA glycosylases, APE1, and pol β, which catalyze the first three steps in BER, proved able to process their substrates without irreversibly disrupting the host nucleosome (Odell et al., 2011). These enzymes all bend DNA, which, in nucleosomes, may require local lifting of DNA substrates away from the histone octamer. Both DNA ligase IIIα (which completes short patch BER) and DNA ligase I (which completes long patch BER and DNA replication) differ in that they fully encircle their DNA substrates (Cotner-Gohara et al., 2010; Pascal et al., 2004). This suggests that binding of either ligase to DNA nicks would require a more extensive disruption of the host nucleosome. Nevertheless, both enzymes proved able to detect and process DNA nicks in nucleosomes (Chafin et al., 2000; Odell et al., 2011). Moreover, LigIIIα-XRCC1, which can bind single base gaps as well as DNA nicks in nucleosomes, enhanced the activity of pol β (which acts just prior to DNA ligase). The enhanced activity of pol β was evident even in the presence of mutations that weaken its binding to XRCC1, which suggested that LigIIIα-XRCC1 renders substrates more readily accessible by altering nucleosome structure (Odell et al., 2011). However, it was not clear what these alterations entail, nor indeed whether host nucleosomes survive the DNA ligation step. Specifically, under conditions of native gel electrophoresis we observed that the fraction of DNA migrating with intact nucleosomes diminished with increasing concentrations of LigIIIα-XRCC1. This result was consistent with the irreversible disruption of host nucleosomes. We have
now determined that this apparent disruption is reversible. We find that [1] LigIIIα-XRCC1 forms relatively salt resistant ternary complexes with DNA gap- or nick-containing nucleosomes, [2] DNA ligation takes place within the DNA nick-containing ternary complex, and [3] nucleosomes can spontaneously reform once ligation is complete and LigIIIα-XRCC1 dissociates. We relate these findings to discrete steps in the folding of DNA ligase IIIα into its active configuration.

3.3 Results

Spontaneous partial unwrapping of DNA from the histone octamer enables DNA ligase IIIα-XRCC1 to bind to DNA nicks in nucleosomes.

We showed previously that the human bifunctional DNA glycosylases human endonuclease III-like protein 1 (hNTHL1) and endonuclease VIII-like protein (NEIL1) are able to bind and remove oxidized thymine residues from nucleosomes in vitro, and thereby initiate BER (Prasad, Wallace, et al., 2007). These reactions occur within ternary complexes that can be visualized in gel mobility shift assays, and proceed without irreversibly disrupting the host nucleosome, or shifting its position. We subsequently demonstrated the handoff of hNTHL1 products in nucleosomes, to apurinic endonuclease (APE1) and from there, to DNA polymerase β (pol β) and LigIIIα-XRCC1 (Odell et al., 2011). The efficiency with which these BER enzymes act in nucleosomes varies considerably with the helical orientation of substrates with respect to the underlying histone octamer (reviewed in (Odell et al., 2013)). This variability reflects two different modes of substrate binding (Maher et al., 2013; Prasad, Wallace, et al., 2007). The first is
direct binding of BER enzymes to substrates that are located at accessible sites in the nucleosome. The second, much less efficient binding mode occurs when spontaneous, partial unwrapping of DNA from the histone octamer (hereafter referred to as "nucleosome unwrapping") exposes substrates to solvent. The extent of nucleosome unwrapping is variable, and the duration of the unwrapped state is on the order of milliseconds (Anderson & Widom, 2000; Li et al., 2005; Li & Widom, 2004; Polach & Widom, 1995; Tims et al., 2011). As a result, the frequency of nucleosome unwrapping events is commonly rate-limiting for processing of occluded substrates in nucleosomes (Maher et al., 2013).

Unlike the helical dependency exhibited by the first three enzymes that act in BER, the efficiency of ligation by LigIIIα-XRCC1 was unaffected by the orientation of DNA nicks in nucleosomes (Odell et al., 2011). The simplest explanation for this finding is that LigIIIα-XRCC1 can only bind DNA nicks when they transiently exposed to solvent by nucleosome unwrapping. If this inference is correct, the efficiency of ligation of a DNA nick located 46 nucleotides from the dyad axis of the nucleosome would exceed that for a nick located only 26 nucleotides from the dyad. Figure 1A depicts the nucleosomes we used test this prediction, while Figure 1B shows that the efficiency of nucleosome assembly was nearly 100%. The sequencing gel in Figure 1C shows DNA ligation products generated by the addition of 15 nM LigIIIα-XRCC1 to 1 nM of the -46 nick-containing nucleosomes or naked DNA. We deliberately positioned these nicks so that they would face into the histone octamer, making them difficult or impossible to detect except when exposed by DNA unwrapping. As Figure 1D indicates, the ligation
efficiency reached 30-40% of that seen in a parallel reaction with naked DNA. In contrast, the ligation efficiency for a similarly oriented nick 20 base pairs closer to the dyad axis was another ~3-fold lower (Figure 1E). This result indicates that nucleosome unwrapping is necessary for ligation of nicks in nucleosomes. A related finding is that the ligation of nicks in nucleosomes exhibits biphasic kinetics (compare graphs in Figures 1D and 1E). As discussed in (Maher et al., 2013), these biphasic kinetics are consistent with a dynamic equilibrium between wrapped and unwrapped nucleosome configurations: when added in molar excess, LigIIIα-XRCC1 will bind and quickly process nick-containing nucleosomes that happen to be in an unwrapped configuration. Ligation of additional nicks can occur only when they are exposed by de novo unwrapping events. This occurs so infrequently that the ligation reaction cannot go to completion within the 30 min time frame of the experiments in Figure 1.

**LigIIIα-XRCC1 does not irreversibly disrupt nucleosomes.**

The nucleosome unwrapping needed for productive binding of LigIIIα-XRCC1 to DNA nicks in nucleosomes requires disruption of some histone-DNA contacts, but does not necessarily mean that nucleosomes are fully disrupted by DNA ligation, any more so than they are by transcription (Kulaeva et al., 2013). To assess the extent to which LigIIIα-XRCC1 binding disrupts histone-DNA contacts, we monitored the accessibility of an Age I restriction site that is normally almost completely resistant to cleavage in the nucleosome (Prasad, Wallace, et al., 2007). This restriction site and the DNA nicks used to monitor ligation activity are located on opposite side of the dyad axis (Figure 1A) but
are relatively close to one another in the nucleosome (Figure 2A). Figure 2B shows ligation of DNA in the 46-nick nucleosomes increasing to ~25% with increasing amounts of LigIIIα-XRCC1; the addition of Age I did not affect ligation efficiency in nucleosomes (Figure 2C) or in naked DNA (data not shown). Likewise, the presence of LigIIIα-XRCC1 did not affect Age I cleavage of naked DNA. In the nucleosome, the addition of increasing amounts of LigIIIα-XRCC1 decreased the extent of Age I cleavage, possibly because LigIIIα-XRCC1 blocked access to the Age I restriction site. This result suggested that the extent to which LigIIIα-XRCC1 binding disrupts histone-DNA contacts is limited, and likely not enough to destabilize the entire nucleosome. We next asked if LigIIIα-XRCC1 affected access to a Psi I site, located 11-16 nt from the nick at position -46 (see schematics in Figures 1A & 2A), and largely occluded in the intact nucleosome (Prasad, Wallace, et al., 2007). Control experiments indicated that LigIIIα-XRCC1 did not block access to Psi I site in naked DNA, and Figure 2D shows that it had little or no effect on Psi I cleavage in the nucleosome. This result suggested that histone-DNA contacts disrupted by the binding of ligase IIIα do not extend much beyond the DNA residues that interact directly with the protein (see Supplemental Figure S4).

We previously reported that the fraction of DNA migrating with intact nucleosomes diminished with increasing concentrations of LigIIIα-XRCC1. This result was consistent with the irreversible disruption of host nucleosomes but did not rule out the possibility that persistent binding of LigIIIα-XRCC1 prevented nucleosome reformation. To examine this possibility, we mixed -46-nick nucleosomes with varying amounts of LigIIIα-XRCC1 for 30 seconds at 37°C, and then quenched the reactions by
the addition of our standard, EDTA-containing native gel loading buffer. The samples were immediately loaded onto native gels, to separate and visualize nucleoprotein complexes (Figure 3A, lanes 3-6). For a parallel set of samples, we quenched reactions with gel loading buffer that supplemented with 1.0 µg/gel lane of naked DNA isolated from Micrococcal nuclease-treated chromatin (see Experimental Procedures). These samples were also immediately loaded onto, and fractionated through, native gels (Figure 3A, lanes 7-10). From a third set of samples, we prepared and fractionated DNA on sequencing gels, to monitor the ligation reactions (Figure 3B). Lanes 4-6 in the gel in Figure 3A show that addition of LigIIIα-XRCC1 produces a super-shifted complex, with the concurrent loss of the intact (nick-containing) nucleosome band; as well, some material is trapped in the well of the gel. Lanes 8-10 in the gel in Figure 3A demonstrate that intact nucleosomes can be recovered from these slow-migrating species, by providing excess naked DNA (i.e. alternate binding sites for LigIIIα-XRCC1). Figure 3B shows the accumulation of ligation products within the 30-second reaction time (see Figure 4 for a comparison with longer incubation times). Figure 3C demonstrates that DNA gap-containing nucleosomes also survive the addition of LigIIIα-XRCC1. Thus, the previously observed enhancement of pol β by LigIIIα-XRCC1 (Odell et al., 2011) is not due to irreversible nucleosome disruption.

LigIIIα-XRCC1 bound not only to nucleosomes containing DNA gaps or nicks, but also to nucleosomes assembled with intact, undamaged DNA (Supplemental Figure S2A). The presence of excess unlabeled nucleosomes in these experiments rules out possible interactions between LigIIIα-XRCC1 and the histone octamer. Experiments
described in the Supplemental Figures S2 & S3 suggest instead that LigIIIα-XRCC1 can bind to the blunt DNA ends in our model nucleosomes, provided those ends contain a 5'-phosphate. This DNA end binding proved to be weaker and more salt-sensitive than the interactions between LigIIIα-XRCC1 and gap- and nick-containing nucleosomes, but is consistent with evidence that DNA ligase IIIα can catalyze intermolecular end-joining (Chen et al., 2009; Kukshal et al., 2015; Taylor et al., 2000). The substrate concentrations used in this current study favor nick ligation over blunt end ligation, and we observed no evidence of the later (Supplemental Figure S1). Nor did we observe ligation of Age I-generated, sticky DNA ends in the above-described assays. These observations indicate that DNA end-binding by LigIIIα-XRCC1 did not compromise the integrity of this study.

**Ligation occurs within the ternary complex that forms upon addition of LigIIIα-XRCC1 to nick-containing nucleosomes.**

The fact that LigIIIα-XRCC1 forms a complex with nick-containing nucleosomes that is reversible suggested that, as with the ternary complexes that form during earlier steps in BER, it is a *bona fide* BER intermediate. If this is true, the ternary structures should contain both substrate and product DNAs. To test this prediction, we incubated nick-46 nucleosomes with 15 nM LigIIIα-XRCC1 for 3 and 30 minutes, and then separated intact nucleosomes from LigIIIα-XRCC1-containing ternary complexes, as before, producing the first-dimension gel shown in Figure 4A. Inspection of the “low mobility” band in Figure 4A shows that its intensity diminishes with reaction time, while the intensity of the band corresponding to the intact nucleosome increased (as quantified
in Figure 4B). This suggested that the LigIIIα-XRCC1-containing ternary complex resolves following ligation. We next excised lanes from a gel identical to that shown in Figure 4A, and fractionated DNA from both the nucleosome and ternary complex, by electrophoresis on a second-dimension sequencing gel (see Experimental Procedures). Figure 4C shows that the ligated DNA products (boxed) are present in both the ternary complex and the intact nucleosome. This result provides further evidence that ligation does not result in nucleosome destruction, and supports a mechanism in which LigIIIα-XRCC1 binds and partially disrupts DNA nick-containing nucleosomes and, after nick closure, dissociates, enabling nucleosomes to reform. The aforementioned decrease in the ternary complex with reaction time is also evident in the two-dimensional gel (as quantified in Figure 4D).

**Linker DNA does not enhance ligation of nicks in nucleosomes by LigIIIα-XRCC1.**

In some instances, linker DNA appears to provide enzymes with leverage or a “running-start” for action on nucleosomes. For example, the efficiency with which human DNA ligase I sealed a DNA nick at the dyad axis increased ~10^4-fold when nucleosomes were assembled with 218 bp 5S rDNA instead of 154 bp DNA (~71 bp vs. ~7 bp total linker DNA) (Chafin et al., 2000). To determine if linker DNA similarly enhances the activity of LigIIIα-XRCC1, we compared its activity toward nick containing nucleosomes assembled with 159 bp DNA to its activity toward a 175 bp nucleosome, which contained an additional 16 additional bp linker DNA proximal to the -46 nick (Figure 5A). The nick-proximal linker was thus ~10 bp longer than the DNA footprint.
associated with DNA ligase IIIβ (an alternative splicing variant of ligase IIIα that does not bind XRCC1; (Cotner-Gohara et al., 2010) and Supplemental Figure S4). As Figure 5B indicates, LigIIIα-XRCC1 functioned equally well on both the short and long linker substrates.

Highly stable, 601 nucleosomes are only marginally permissive for DNA ligation by LigIIIα-XRCC1.

To assess the impact of nucleosome stability on the capacity of LigIIIα-XRCC1 to drive nucleosomes into a ligation-permissive configuration, we compared the efficiency of ligation of nicks in nucleosomes containing 5S rDNA to those containing the synthetic “601” DNA (see Experimental Procedures). Competitive reconstitution and thermal denaturation studies (Maskell et al., 2015; Thastrom, Lowary, et al., 2004) indicate that 601 nucleosomes are much more stable than 5S rDNA-containing nucleosomes. Their exceptional stability makes 601 nucleosomes relatively tractable for in vitro studies, but they may not as faithfully reflect chromatin dynamics in cells, for as Widom and colleagues observed in (Thastrom et al., 1999), "even the highest-affinity sequence regions of eukaryotic genomes are not evolved for the highest affinity of nucleosome positioning power." With this caveat in mind, we treated nick-containing 601 nucleosomes containing variable lengths of linker DNA (Figure 6A). The overall ligation efficiency was three- to four-fold lower than we observed for the 5S rDNA nucleosomes (Figure 6B). As well, the addition of variable amounts of linker DNA to 601 DNA-nucleosomes had no impact on either ligation efficiency or ternary-complex formation.
The octamer structure within 601 nucleosomes is virtually identical to that in other nucleosomes (Vasudevan et al., 2010). This suggests that the structural elements that reduce ligation efficiency in 601 nucleosomes reside in the exceptional rigidity of the nucleosome that 601 DNA forms. This could manifest itself at the level of nucleosome unwrapping that exposes DNA nicks or the DNA untwisting that accompanies the folding of ligase IIIα into its active configuration.

Ligation of nicks in nucleosomes by ligase IIIα does not require XRCC1.

XRCC1 serves not only as a scaffold for the binding of downstream BER enzymes but also binds preferentially to DNA nicks and single base gaps, with reported K_D's of ~65 nM and ~34 nM, respectively (Mani et al., 2004). The nick and gap binding affinities for DNA ligase IIIβ are probably similar in magnitude although reported K_D values for a ligatable nick substrate range between ~5 and 200 nM at physiological salt concentrations, depending on assay methods (Cotner-Gohara et al., 2008; Leppard et al., 2003). The relatively robust activity that we observed in reactions containing only 15 nM LigIIIα-XRCC1 is consistent with a K_D value at the lower end of the range. However, it is possible that the overall DNA binding affinity for LigIIIα-XRCC1 exceeds that of either subunit alone, as appears to be the case for complexes between pol β and XRCC1 (Marintchev et al., 1999). To assess the extent to which XRCC1 contributes to binding of DNA nicks in nucleosomes, we compared LigIIIα-XRCC1 to LigIIIα alone, in reactions with 5S rDNA-based, nick-46 nucleosomes. Figure 7A shows that DNA ligase IIIα alone exhibited significant activity on nicks in nucleosomes. Although the extent of
nucleosome ligation was somewhat less than we observed for LigIIIα-XRCC1 (Figure 7B), pre-incubating LigIIIα for several minutes on ice with E. coli-source XRCC1 had no impact on ligation activity (data not shown). Moreover, XRCC1 alone did not detectably interact with gap-containing naked or nucleosomal DNA in a gel mobility shift assay (Figure 7C). These observations suggest that XRCC1 does not play a vital role in the discovery or binding of LigIIIα to nicks in nucleosomes. It is possible, however, that the impact of XRCC1 will become evident only in the presence of certain posttranslational modifications or other binding partners.

3.4 Discussion

Prior to this study we had determined that the entire four-step BER reaction can be reconstituted with nucleosomes containing oxidized bases, and that the three enzymes that catalyze the first three steps in BER each form ternary complexes with the host nucleosomes (Odell et al., 2011). We had also determined that host nucleosomes are not irrevocably disrupted by the action of these enzymes. The evidence presented in this study indicates that host nucleosomes also survive LigIIIα-XRCC1, which catalyzes the fourth and final step in BER. In an effort to understand how BER enzymes can detect and process substrates in nucleosomes, we have developed heuristic models that incorporate available data on the structure of nucleosomes and BER enzymes. We assume that enzyme-substrate complexes that form in nucleosomes are similar in structure to those in naked DNA, and that dynamic properties intrinsic to nucleosomes enable them to accommodate these enzyme-substrate complexes. For example, as noted earlier, several
DNA glycosylases as well as APE1 and pol β can bind and efficiently process optimally oriented substrates in nucleosomes. None of these enzymes fully encircle DNA, which may explain their capacity for direct binding. However, all three enzymes bend DNA as they fold into their active configurations, which may require that local DNA lifted away from histone octamer. Local DNA untwisting could in principle provide the longer DNA path that this lifting requires.

Unlike the first three enzymes that act in BER, DNA ligase IIIα fully encircles its substrate in its active configuration, and appears unable to bind nucleosomes without nucleosome unwrapping (or an equivalent conformational change) that disrupts histone-DNA contacts. The model depicted in Figure 8 depicts the multiple discrete steps that the ligation of DNA nicks in nucleosomes by LigIIIα-XRCC1 likely entails. Most DNA-binding proteins locate targets in genomic DNA in two stages, beginning with a diffusion-driven search for DNA in three-dimensional space. Once the protein encounters DNA, it binds in a non-specific fashion, which enables it to engage in a much more efficient (but still diffusion-driven), “one-dimensional” search along DNA. This search ends when the protein either dissociates or encounters its target, which it binds in a more specific fashion (von Hippel & Berg, 1989). Non-specific DNA binding commonly depends heavily on electrostatic interactions, making it salt-sensitive. This has been observed for ligase IIIα (Cotner-Gohara et al., 2008), which suggests that during the search for DNA nicks, ligase IIIα binds DNA in a non-specific configuration, as depicted Figure 8B. Although LigIIIα-XRCC1 is depicted as binding to linker DNA, it may also associate non-specifically with DNA in nucleosomes. In its unliganded configuration and
in naked or linker DNA, ligase IIIα-XRCC1 might track along the sugar-phosphate backbone of DNA, as is the case for the DNA glycosylases that initiate BER (Dunn et al., 2011; Nelson et al., 2014; Qi et al., 2012). On the other hand, if the initial encounter is with histone octamer-associated DNA, LigIIIα-XRCC1 may be forced to engage in a less efficient search mode that entails 'hopping' (in N x ~10 bp intervals) along one face of the DNA helix. The alternative (i.e. tracking along the sugar-phosphate backbone) would require either the transient displacement of DNA from the histone octamer or the rotation of DNA to relative to the underlying histone octamer. Both alternatives would require breaking multiple discrete interactions between DNA and the histone octamer, and the latter process would also force DNA to bend in non-preferred directions. Although this probably occurs for factors whose elongation is driven by nucleotide hydrolysis (e.g. RNA polymerase), it is unlikely to be energetically feasible for diffusion-driven search processes.

The periodic partial unwrapping of nucleosomal DNA that exposes a DNA nick to solvent, depicted in Figure 8C, is likely to be rate-limiting in the discovery and processing of DNA nicks in chromatin. DNA ligase IIIα contains a poly-ADP ribose polymerase 1 (PARP1)-like zinc finger (ZnF in Figure 8A), which will bind to DNA gaps or nicks (Mackey et al., 1999). This likely triggers its folding into a more specific binding configuration (Figure 8D). The "jackknife" model in (Cotner-Gohara et al., 2010) proposes that the zinc finger folds toward a nearby DNA binding domain (DBD), to form DNA binding “module #1”. Then, if ligase IIIα is bound to a ligatable nick, the NTase (nucleotidyltransferase) and OBD (oligonucleotide/oligosaccharide-binding domain)
domains fold toward one other, and then toward the DNA binding domain, forming DNA
binding “module #2.” With the “transfer” of the DNA binding domain from binding
module #1 into binding module #2, the zinc finger is likely displaced, allowing the
substrate to be passed to module #2, which contains the enzyme's catalytic core. This
step, which is depicted in Figure 8E, culminates with ligation. Figure 4 demonstrates that
ligation occurs within a nucleosome-containing complex, and that this complex resolves
itself over time, inversely correlating with the accumulation of ligated substrate. This
suggests that, once ligation is complete, ligase IIIα returns to its non-specific DNA
binding configuration (Figure 8F). This would likely promote enzyme turnover and
permit nucleosomes to refold (Figure 8G).

In the event that ligase IIIα binds to a single base DNA gap, it may not fold to
“completion.” Nevertheless, the folding to form DNA binding module #1 is apparently
sufficient to render its association with DNA more salt tolerant. In chromatin, this may
increase the dwell time of LigIIIα-XRCC1, and inhibit DNA rewrapping. This could
explain how LigIIIα-XRCC1 facilitates the activity of pol β (Figure 8H; refer to (Odell et
al., 2011)), although we have yet to rule out the alternative possibility that pol β is
stimulated by the gap binding activity of XRCC1 (Mani et al., 2004). In our in vitro
model system, XRCC1 had little or no impact on the efficiency of ligation.

Summary and perspective.

We have shown that dynamic properties intrinsic to nucleosomes facilitate the
ligation of nicked DNA ligation by purified human LigIIIα-XRCC1, and help explain the
survival of host nucleosomes. That the final ligation step of BER by ligase IIIα can occur on nucleosomes raises interesting parallels with the replicative ligase, DNA ligase I. The packaging of newly replicated DNA into nucleosomes is coordinated with movement of the replication fork through physical interactions between replication machinery and nucleosome assembly factors ([Smith & Stillman, 1991]; reviewed in (Krude, 1999)). Thus, the ligase I-mediated sealing of DNA nicks during the processing of Okazaki fragments may occur after DNA has been incorporated into nucleosomes (Bielinsky & Gerbi, 1999), and, as described earlier, ligase I is able to seal nicks in nucleosome in vitro (Chafin et al., 2000). That both ligases can act on nucleosome substrates is consistent with reports that ligase I is able to complete BER in the absence of ligase III (Gao et al., 2011; Simsek et al., 2011), and that ligase III can function as a replicative ligase (Arakawa & Iliakis, 2015). Replication-coupled nucleosome assembly requires specific histone secondary modifications and histone chaperones that might be present during the replicative process, and assist in ligation of Okazaki fragments. It is possible as well that chromatin-remodeling agents associated with newly replicated chromatin promote or amplify the unwrapping events that enable ligase I to load onto nucleosomal DNA. As with the DNA ligases, RNA polymerase II encircles DNA, but it differs in its capacity to “ratchet through” an entire nucleosome in the absence of ATP-dependent chromatin remodelers (Hodges et al., 2009). Its transit through a nucleosome requires that it (or DNA) rotate relative to the histone octamer, thus disrupting DNA-histone contacts along the way (Kulaeva et al., 2013). This transit appears to entail loss of a single H2A/H2B dimer from the host nucleosome, but does not alter translational position of the
nucleosome (Kireeva et al., 2002). Based on our gel mobility shift assay, ligation at the nick positions tested did not result in histone loss. However, it is possible that the integrity of the octamer (and the capacity of the host nucleosome to refold following ligation) varies with the site of nick repair. These many considerations and experimental variables not yet explored suggest that we are far from being able to account for how BER occurs in chromatin in vivo. Further progress will require investigating the possible contributions to BER of histone modifiers and chaperones, and chromatin remodeling agents, as well as post-translational modifications that regulate BER enzymes.

3.5 Experimental Procedures

Preparation of DNA substrates.

We assembled undamaged nucleosomes and DNA nick- or gap-containing nucleosomes with either the synthetic “601” nucleosome positioning sequence (Lowary & Widom, 1998) or the Lytechinus variegatus 5S rDNA (“5S”) nucleosome positioning sequence (Simpson & Stafford, 1983). 601 DNA adopts a single helical orientation and translational position with respect to the underlying histone octamer, allowing us to predict the location and orientation of each base in the sequence. 5S rDNA also adopts a single helical orientation but, depending on reconstitution conditions, may adopt one or a few translational positions, separated by ~10 bp intervals (Dong et al., 1990; Flaus et al., 1996; Prasad, Wallace, et al., 2007). To prepare DNA for these nucleosomes, we annealed equimolar amounts of a top-strand DNA oligomer containing sequences 3’ to the desired nick or gap to a full-length (147-183 nt) bottom strand oligomer. (All
oligomers were purchased from Integrated DNA Technologies.) The top-strand oligomer contained a 5' phosphate moiety, to permit ligation, and was lightly end-labeled with $\gamma$-$^{32}$P-ATP (Perkin Elmer) and polynucleotide kinase (New England Biolabs), to help track DNA through subsequent steps. The resulting duplex was mixed with the thermostable DeepVent exo- polymerase (New England Biolabs), and extended in a linear PCR amplification reaction (30 cycles X 30 sec at 95°C, 30 sec at 46°C, and 30 sec at 68°C). To verify quantitative extension, reaction products were examined by denaturing PAGE, and then annealed to a $\gamma$-$^{32}$P-ATP-end-labeled oligomer containing sequences 5' to the desired nick or gap. We were able to use this second $^{32}$P labeled oligomer to monitor ligation because it migrated differently from the first on gels, and was labeled to a higher specific activity. To assemble undamaged DNA substrates, we omitted the 3’ oligomer annealing and extension step, and instead annealed and extended the 5’ oligomer.

**Nucleosome reconstitutions.**

The $^{32}$P-labeled DNAs were combined with an 80-fold molar excess of unlabeled carrier DNA, purified from chicken chromatin that had been digested with micrococcal nuclease to mono- di- and tri-nucleosome-lengths, as described in (Prasad, Wallace, et al., 2007). This nuclease cleaves 3’ of the phosphodiester backbone in DNA, producing DNA fragments with 5’-hydroxyl and 3’-phosphate groups. Hence, this DNA did not interfere with our ligation assays (see Supplemental Figure S1). To this DNA mixture, we added a slight molar excess of histone octamer in 25 mM HEPES, pH 8.0, 1 mM EDTA, 1 mM dithiothreitol and 2 M sodium chloride (reconstituted from *E. coli*-expressed
recombinant human or *Xenopus* histones, and purified as described in (Dyer et al., 2004; Luger et al., 1999)). Samples were then slowly dialyzed, in a step-wise fashion, into the same buffer lacking sodium chloride, beginning at 37°C for the first step, and ending with an overnight dialysis at 4°C (note that nucleosomes are stable in low salt while the histone octamer is not). Reconstitution efficiencies were assessed using 5% native PAGE in 1/2 X Tris-Borate-EDTA (electrophoretic mobility shift assay or “EMSA” gels), and were virtually quantitative for the 601 nucleosomes and typically >90% for 5S rDNA nucleosomes. The end-labeled nucleosomes were adjusted to 2.5 nM (~ 200 nM total nucleosome concentration) for storage at 4°C, and diluted to 1 nM (80 nM total nucleosome concentration) just prior to use in ligation assays. This is well above the concentration where one can see dilution-driven nucleosome dissociation (Pederson et al., 1986).

**Ligation assays.**

Human DNA ligase IIIα-XRCC1 was expressed in and purified from baculovirus-infected Sf9 insect cells. *E. coli*-expressed Human LigIIIα was from Enzymax, LLC, while *E. coli*-expressed XRCC1 was a gift from Michael Weinfeld. Unless otherwise noted, ligation reactions contained 1 nM of 32P-end-labeled nucleosomes or naked DNA, lacking or containing a single DNA nick or gap, and ~80 nM unlabeled nucleosomes or naked DNA, in our standard assay buffer (25 mM HEPES, pH 8.0, 100 mM sodium chloride, 5 mM magnesium chloride, 1 mM dithiothreitol, 1 mM ATP, and 0.1 mg/mL bovine serum albumin), in a final volume of 5 to 10 µL. Reaction mixtures were
incubated at 37°C with the indicated enzyme(s), at concentrations and times specified in the text or figure legends. To monitor ligation, reactions were quenched by adding 1/4th volume of 0.5% sodium dodecyl sulfate, 50 mM EDTA, and 0.5 mg/mL of freshly added proteinase K. Samples were incubated for 30 minutes at 37°C, then mixed with an equal volume of formamide containing 20 mM EDTA and 2 mg/ml each of bromophenol blue and xylene cyanol, and fractionated on 8% denaturing gels. Gels were dried, exposed to phosphoimaging screens (BioRAD), and bands of interest quantified using a PharosFX Plus phosphoimager with Quantity One software (BioRAD). We calculated ligation as a percentage of the total DNA substrate, after adjusting for background (using the global subtraction method). The impact of small amounts of contaminating naked DNA in the nucleosome reactions was computationally eliminated, as described in (Prasad, Wallace, et al., 2007).

**Restriction enzyme assays for nucleosome integrity.**

To assess the fate of nucleosomes during ligation, we monitored the availability of restriction endonuclease sites in 5S rDNA that are normally inaccessible (Odell et al., 2010; Prasad, Wallace, et al., 2007). We added 25 units of Age I (New England Biolabs) to 5µL reaction mixtures containing 0 to 15 nM LigIIIα-XRCC1, 1 nM of -46 nicked DNA or nucleosomes, and 80 nM unlabeled naked DNA or nucleosomes. After 45 min at 37°C, we measured the extent of ligation by quenching a portion of each reaction, and processing samples as described above. To measure the extent of Age I cleavage in the presence or absence of LigIIIα-XRCC1, we added to a second portion of each reaction an
equal volume of 100 mM Tris base, 25 mM EDTA, 0.2% sodium dodecyl sulfate, 2 mg/ml bromophenol blue and 12% glycerol, and fractionated samples through native, 8% gels. We monitored access to a Psi I site closer to the DNA nick, in an identical manner, using 15 nM LigIIIα-XRCC1 and 2.5 U of Psi I.

_Ternary complex analyses_ - To visualize the impact of adding LigIIIα-XRCC1 to intact or nick- or gap-containing nucleosomes, portions of the above-described reactions were mixed with equal volumes of 10% glycerol, 25 nM HEPES, pH 8.0, and 20 mM EDTA, and then immediately fractionated through 5% EMSA gels, and visualized as before. To determine if DNA ligation occurs within the ternary complex that LigIIIα-XRCC1 forms with nick-containing nucleosomes, ternary complexes were separated from nucleosomes, using EMSA gels. Gel lanes were then excised and incubated first in 0.1% sodium dodecyl sulfate, 10 mM EDTA (pH 8.0), and 0.1 mg/mL of proteinase K for 30 minutes at 50°C, to degrade protein, and then in 0.1 M sodium hydroxide, 5% glycerol, and 0.0125% each of bromophenol blue and xylene cyanol for 10 minutes, to denature DNA. DNA associated with individual particles was then fractionated through an 8% sequencing gel.

**Acknowledgements:** We thank Drs. Joyce Heckman and Robyn Maher for critical comments on the manuscript, Dr. Michael Weinfeld for purified XRCC1, Dr. Amalthiya Prasad for purified chicken nuclei, and April Averill for the purification of histones.
**Conflict of interest:** The authors declare that they have no conflicts of interest with the contents of this article.

**Author contributions:** WJC conducted the experiments, analyzed the results, prepared all the figures, and wrote the paper. DSP conceived the idea for the project, helped analyze the results, and wrote the paper with WJC. SSW and AET helped to write the paper.
3.6 References


### 3.7 Footnotes

**Funding:** This study was funded by grants from the National Cancer Institute [P01-CA098993 to S.S.W. and CA92584 to A.E.T.], and the National Institute of Environmental Health Sciences [ES012512 to A.E.T.]. The authors of this study have no financial conflicts of interest.

**Abbreviations:** The abbreviations used are: 5'-2-deoxyribose phosphate, dRP; 5S or 5S rDNA, *Lytechinus variegatus* 5S rDNA nucleosome positioning sequence; APE1, apurinic endonuclease; BER, base excision repair; bp, base pair; EMSA, electrophoretic mobility shift assay; NEIL1, endonuclease VIII-like protein; hNTHL1, human
endonuclease III-like protein 1; LigIIIα, DNA ligase IIIα; nt, nucleotide; NTase, nucleotidyltransferase; OB-fold, oligonucleotide/oligosaccharide-binding fold; PARP1, poly-ADP ribose polymerase 1; pol β, DNA polymerase β; XRCC1, X-ray repair cross-complementing protein 1;

3.8 Figure Legends

FIGURE 1. Spontaneous partial unwrapping of DNA from the histone octamer enables DNA ligase IIIα-XRCC1 to bind to DNA nicks in nucleosomes. (A) Schematic depicting relevant features of the nucleosomes used in the experiments shown in Figures 1-3. The grey oval represents the predominant translational position of the 5S rDNA nucleosome; minor translational variants, in which the octamer is positioned ~10 bp to either the left or right, also occur (Maher et al., 2013; Prasad, Wallace, et al., 2007; Simpson & Stafford, 1983), but did not affect the interpretation of results. Ligatable, single strand DNA nicks were positioned either -46 or -26 nt away from the nucleosome center (the dyad axis). Also shown are restriction endonuclease cleavage sites used (in Figure 2) to monitor nucleosome integrity during ligation. (B) Representative nucleosome preparation examined by native PAGE. (C) Representative 8% sequencing gel used to monitor ligation of nicks in nucleosomes and naked DNA controls, as detailed in the Experimental Procedures. Ligation products migrate as 183 nt bands. The minor band denoted by an asterisk corresponds to unligated DNA located 3’ to the nick, and does not interfere with the quantification of ligation extent. (D) LigIIIα-XRCC1 ligated a significant fraction of the DNA nicks at position -46, even though these nicks faced into
the histone octamer, making them accessible only during episodes of spontaneous partial unwrapping of DNA from the histone octamer (see text for further discussion). (E) Ligation efficiency dropped a further ~3-fold for a nick positioned 20 nt further in from the nucleosome edge (position -26). This supports the hypothesis that nucleosome unwrapping is rate-limiting LigIIIα-XRCC1. Error bars represent standard deviations.

FIGURE 2. Ligation by LigIIIα-XRCC1 does not irreversibly disrupt nick-containing nucleosomes. (A) Schematic depicting nick location with respect to the restriction sites examined in the context of nucleosome wrapping. (B) Ligation of a DNA nick 46 nt from the dyad axis does not interfere with nucleosome-mediated protection from cleavage at an Age I site. 1 nM nick-46 nucleosomes were incubated with 25 Units Age I together with 1 to 15 nM LigIIIα-XRCC1, and ligation and cleavage amounts determined as described in the Experimental Procedures. Age I cleavage was calculated as a fraction of naked DNA, while nucleosome ligation was reported as an absolute fraction. (C) The presence of Age I in the ligation reactions reported in Figure 2B had no significant impact on the efficiency of ligation of the -46 nicks in nucleosomes, or in naked nick-46 DNA controls (where ligation reached ~77%; not shown). (D) The left-side axis shows that 15 nM LigIIIα-XRCC1 did not appreciably affect cleavage at a Psi I located approximately midway between the -46 DNA nick and the edge of the nucleosome (Figures 1A & 2A). The right-side axis shows efficient ligation of nick-46 nucleosomes in the absence Psi I. Ligation also occurred in the presence of Psi I, but it
was not possible to quantify the extent of ligation because the Psi I site lies between the DNA nick and the 5’-^{32}P label).

**FIGURE 3. Binding of LigIIIα-XRCC1 does not irreversibly disrupt nucleosomes.** (A) Lanes 1 and 2 show migration in native (mobility-shift) gels of an end-labeled, 183 bp DNA fragment containing a single DNA nick, in the absence and presence of LigIIIα-XRCC1. Lanes 3-6 show migration of nucleosomes assembled with the same DNA, and then incubated in the absence and presence of 1, 15, or 50 nM LigIIIα-XRCC1. In all samples containing LigIIIα-XRCC1, substrate concentrations were 1 nM, and incubation times were 30 sec. Ligation reactions were halted by addition of our standard, EDTA-containing gel-loading buffer (lanes 3-6), and then run on mobility-shift gels. Samples in lanes 7-10 were treated identically to those in lanes 3-6, except for the inclusion of 1 µg of non-specific carrier DNA in the gel-loading buffer. Note that addition of nonspecific DNA to nucleosome-LigIIIα-XRCC1 reaction mixtures enables nucleosomes to reform, indicating that all nucleosome components are present in the low mobility complexes in lanes 5-6. (B) DNA from LigIIIα-XRCC1 reactions was fractionated through denaturing gels to quantify ligation products generated in the 30 sec. reactions. Error bars represent standard deviations. (C) Mobility shift gels of 1 nM -46 gap-containing nucleosomes incubated with varying amounts of LigIIIα-XRCC1, as in panel A.

**FIGURE 4. Ligation occurs within the ternary complex that forms upon addition of LigIIIα-XRCC1 to nick-containing nucleosomes.** (A) 1 nM -46 nick nucleosomes
incubated without or with 15 nM LigIIIα-XRCC1 for 3 and 30 minutes were fractionated through mobility shift gels. (B) Quantification of bands in gels such as those in (A) shows a decline in the nucleosome-LigIIIα-XRCC1 ternary complex, and a corresponding increase in intact nucleosome with time. This suggests that the ternary complex decays once ligation is complete. (C) Lanes were cut out of a mobility shift gel identical to (A), and DNA in each of the complexes was fractionated through a second dimension sequencing gel (see Experimental Procedures). D=DNA, N=Nucleosome, S=super-shifted complex. The boxes indicate nicked and ligated DNAs that are present within both the nucleosome band and the ternary complex. This result indicates that productive ligation occurs within low mobility ternary complex. (D) The fraction of total ligation product in the low mobility ternary complex declines with time, further supporting the inference that, once ligation is complete, the ternary complex decays, allowing the host nucleosome to return to its native configuration. Error bars represent standard deviations (n=3).

FIGURE 5. Ligation of nicks in nucleosomes by LigIIIα-XRCC1 is not enhanced by linker DNA. (A) Schematic depicting 5S *L. variegatus*-based nick-46 nucleosomes flanked by short DNA segments (159 bp total), or with another 16 bp of linker DNA added to the nick-proximal (“left”) side of the same nucleosome (175 bp total). (B) Ligation products generated in 30 min by 15 nM LigIIIα-XRCC1 at 37°C. There is no significant difference in ligation efficiency between the short and left-overhang nucleosomes. Error bars represent standard deviations (n=3).
FIGURE 6. Reduced ligation efficiency in highly stable 601 nucleosomes. (A) DNA nick-containing nucleosomes, containing either no linker DNA (“short”) or 26 bp linker DNA added to the nick-proximal side of the nucleosome (“left overhang”) were incubated with 15 nM LigIIα-XRCC1 at 37°C for 30 min, and the extent of ligation measured as before. As the histogram indicates, very little ligation occurred in 601 nucleosomes, in reaction conditions identical to those used to investigate the impact of linkers on ligation of nicks in 5S rDNA-containing nucleosomes (Figure 5B). (B) Ligation efficiency in the 601-nucleosomes compared to that of the 5S rDNA-nucleosomes. Values shown are based on data from Figures 6A and 5B, after normalizing to the extent of ligation obtained in reactions with the corresponding naked DNAs. The extent of ligation of nicks in 5S rDNA nucleosomes was invariably 3-4 fold higher than in the 601 nucleosomes, regardless of the presence or absence of linker DNA. (C) The efficiency of ligation of nicks in 601 nucleosomes is not limited by the non-specific binding capacity of LigIIα-XRCC1. 601 nucleosomes containing intact DNA were incubated with LigIIα-XRCC1, and analyzed in gel shift assays, as in Figure 3. The histogram shows that the fraction of left overhang 601 nucleosomes that formed non-specific complexes with LigIIα-XRCC1 was significantly higher than the efficiency of ligation. This result indicates that non-specific binding likely precedes specific binding, and that unwrapping-mediated exposure of nicks, or an even later step is rate-limiting step for the ligation of nicks in 601 nucleosomes. Error bars represent standard deviations.
**FIGURE 7. Ligation of nicks in nucleosomes by ligase IIIα does not require XRCC1.**

(A) DNA nick-containing, 5S rDNA-based nucleosomes were incubated at 37°C for 30 min with either 15 nM of Sf9-expressed LigIIIα-XRCC1 or an equivalent amount of *E. coli*-expressed LigIIIα (using concentrations specified by the supplier), and the extent of ligation was measured as before. (B) Comparison of nucleosome ligation efficiency between LigIIIα-XRCC1 and LigIIIα. Values shown are based on data from Figure 7A, after normalizing to the extent of ligation obtained in reactions with the corresponding naked DNAs. In the absence of XRCC1, the extent of ligation was ~2/3 that observed with LigIIIα-XRCC1. (C) In gel shift experiments conducted as in Figure 3, *E. coli*-purified recombinant XRCC1 did not detectably alter the mobility of nucleosomes containing a single nucleotide gaps DNA/nucleosomes. This result suggests that both the nucleosome binding and conformation changes leading up to DNA ligation by LigIIIα-XRCC1 are driven primarily by DNA ligase IIIα.

**FIGURE 8. Discrete steps in the recognition and ligation of DNA nicks in nucleosomes, by LigIIIα-XRCC1.** The multi-step model depicted incorporates the results from this study, as well as DNA binding studies by (Cotner-Gohara et al., 2008) and structural studies by (Cotner-Gohara et al., 2010). Step 1 depicts the nonspecific binding of LigIIIα-XRCC1 to DNA in chromatin. Step 2 depicts the exposure of DNA nicks or gaps during periodic, partial unwrapping of DNA from the histone octamer, as discussed in detail in the text. Under the enzyme-excess conditions used in this study,
nucleosome unwrapping is rate limiting for the de novo binding of LigIIIα-XRCC1 to nicks or gaps in nucleosomes (for further discussion, see (Maher et al., 2013)). This may not invariably be the case in complete BER reactions where substrates may be handed off from upstream enzymes. LigIIIα-XRCC1 binding to gapped DNA promotes folding into a more salt-resistant configuration, which may promote or stabilize changes that make nucleosomes more permissive for DNA Pol β (Odell et al., 2011). LigIIIα-XRCC1 binding to nicked DNA may trigger additional enzyme folding and transfer of the nick from binding module 1 to the catalytic site in binding module 2. This folding and (possibly concurrent) changes in nucleosome conformation needed to accommodate DNA LigIIIα may be inhibited in the unusually, stable 601 nucleosome. See text for further discussion. ZnF = zinc finger, DBD = DNA binding domain, NTase = nucleotidyltransferase domain, OBD = OB-fold domain.

SUPPLEMENTAL FIGURE S1. Carrier DNA does not interfere with ligation assays. (A) We obtained carrier DNA by digesting chicken chromatin, isolated from red blood cells as described (Prasad, Wallace, et al., 2007), with enough Micrococcal nuclease to cleave DNA into one to ~three nucleosome-long fragments. These DNA fragments were purified and quantified by standard methods. Here, and throughout this paper, we have expressed the molarity of both carrier and test DNAs in 'single-nucleosome units.' Hence, 500 nM carrier DNA is sufficient to assemble 500 nM nucleosomes. In the experiment shown, we added 0-500 nM naked, carrier DNA to 1 nM of 32P-end-labeled 183 bp DNA, containing a single nick. (The labeled DNA was
identical to that used to assemble 5S rDNA-based nucleosomes for many of the experiments reported in this paper.) After incubating these DNA mixtures with 5 nM LigIIIα-XRCC1, for 30 min at 37°C, we separated substrate and product DNAs on sequencing gels such as those shown in Figure 1B, and quantified the fraction of substrate ligated, using phosphor-imagery. Importantly, we did not detect any dimeric, blunt-end ligation products, which would have migrated at the position denoted by the asterisk. (B) The results graphed indicate that carrier DNA did not suppress ligation, even at concentrations well above the 80 nM carrier used in all assays.

SUPPLEMENTAL FIGURE S2. Interactions between LigIIIα-XRCC1 and undamaged nucleosomes. (A) Gel mobility shift assay of 1 nM of 5S rDNA-based, undamaged nucleosomes, in the absence (lanes 1 and 5) and presence of increasing concentrations of LigIIIα-XRCC1 (lanes 2-4 and 6-8). As in the experiments shown in Figures 3A and 3C, samples in lanes 5-8 were identical to those in lanes 1-4, expect for the addition of unlabeled competitor DNA just prior to loading samples on the gel. The low mobility complex evident in lanes 3-4 is indistinguishable from the complexes formed by the addition of LigIIIα-XRCC1 to nick- and gap-containing nucleosomes. (B) The fraction of nucleosomes shifted into low mobility complexes in all the EMSA experiments conducted (see Figures 3A, 3C, and S2A, respectively for representative gels) was quantified and plotted as function of nucleosome type and LigIIIα-XRCC1 concentration. For ease of comparison, the same data were plotted in the histograms shown in (C). The somewhat lower affinity of LigIIIα-XRCC1 for nucleosomes with
undamaged DNA is apparent only when enzyme concentrations are reduced to 5.0 nM. However, as shown in Supplemental Figure S3, the differences in affinity are amplified by increasing salt concentrations. All of these mobility shift experiments were conducted in the presence of an ~80-fold excess of unlabeled carrier nucleosomes. The fact that these carrier nucleosomes did not abolish the binding of LigIIIα-XRCC1 to end-labeled, undamaged nucleosomes suggests that LigIIIα-XRCC1 can interact with the 5′-phosphate-containing, blunt DNA ends of our end-labeled model nucleosomes but not with the 5′-OH terminated ends of the cold carrier DNA. This inference is consistent with the observation that DNA ligase III can join blunt ended fragments (in the absence of XRCC1; (Taylor et al., 2000)), although the enzyme and substrate concentrations used in this study were not high enough to catalyze intermolecular ligation (Supplemental Figure 1).

**SUPPLEMENTAL FIGURE S3. Preferential suppression of interactions between LigIIIα-XRCC1 and undamaged nucleosomes at elevated salt concentrations.** To further investigate the basis for interactions between LigIIIα-XRCC1 and undamaged nucleosomes, we repeated the LigIIIα-XRCC1 binding studies, using buffers containing enough salt to suppress the reported binding of DNA ligase IIIα to intact double stranded DNA (Cotner-Gohara et al., 2008) but not so high as to abolish ligase activity (Mackey et al., 1999) or destabilize nucleosome substrates. The high salt buffer suppressed binding of LigIIIα-XRCC1 to undamaged, blunt-ended naked DNA to the same extent as for undamaged nucleosomes (green line in panel A), ruling out possible DNA-independent
interactions between LigIIIα-XRCC1 and the histone octamer. The binding of LigIIIα-XRCC1 to nucleosomes containing undamaged DNA, a DNA nick, and a DNA gap, at elevated salt concentrations is reported in panels A, B and C, respectively. For ease of comparison, data from Figure 3C showing the binding in our standard, lower salt buffer are re-graphed here. (D) Bar graph comparing the low and high salt binding of 15 nM LigIIIα-XRCC1 to the three different substrates. While gapped and nicked substrates demonstrate a reduced capacity to shift the nucleosome, only the undamaged substrate demonstrates significant difference between 100 and 230 mM sodium chloride. (E) Graph depicting ligation activity of (A), 30 second end point reaction (data from Figure 3A is re-graphed here for comparison). Error bars represent standard deviations.

SUPPLEMENTAL FIGURE S4. Extensive interactions between LigIIIβ and DNA.

The crystal structure of LigIIIβ bound to a 20 bp nicked DNA substrate (PDB ID 3L2P) was analyzed using NUCPLOT (http://www.ebi.ac.uk/thornton-srv/software/NUCPLOT/) (Luscombe et al., 1997), to identify likely interacting residues. Interactions between LigIIIβ and DNA bases and the nick itself cover only ~3 bp, but there are numerous additional interactions with the sugar-phosphate backbone covering about one and a half helical turns (16 bp). The actual crystal structure (Cotner-Gohara et al., 2010) also reveals significant distortion and partial unwinding in the DNA at the nick location. Therefore, not only does ligase need at least 16 bp of DNA for complete binding, it must also be able to “bend” the DNA into a ligation-permissible confirmation upon binding. These extensive requirements may explain why LigIIIα-XRCC1 can only
bind nicks and gaps in nucleosomes when they are exposed by partial unwrapping of DNA from the histone octamer.
3.9 Figures

Figure 3-1. Spontaneous partial unwrapping of DNA from the histone octamer enables DNA ligase IIIα-XRCC1 to bind to DNA nicks in nucleosomes.
Figure 3-2. Ligation by LigIIIα-XRCC1 does not irreversibly disrupt nick-containing nucleosomes.
Figure 3-3. Binding of LigIIIα-XRCC1 does not irreversibly disrupt nucleosomes.
Figure 3-4. Ligation occurs within the ternary complex that forms upon addition of LigIIIα-XRCC1 to nick-containing nucleosomes.
Figure 3-5. Ligation of nicks in nucleosomes by LigIIIα-XRCC1 is not enhanced by linker DNA.
Figure 3-6. Reduced ligation efficiency in highly stable 601 nucleosomes.
Figure 3-7. Ligation of nicks in nucleosomes by ligase IIIα does not require XRCC1.
Figure 3-8. Discrete steps in the recognition and ligation of DNA nicks in nucleosomes, by LigIIIα-XRCC1.
Supplemental Figure 3-1. Carrier DNA does not interfere with ligation assays.
Supplemental Figure 3-2. Interactions between LigIIIα-XRCC1 and undamaged nucleosomes.
Supplemental Figure 3-3. Preferential suppression of interactions between LigIIIα-XRCC1 and undamaged nucleosomes at elevated salt concentrations.
Supplemental Figure 3-4. Extensive interactions between LigIIIβ and DNA.
CHAPTER 4: CONCLUSIONS AND FUTURE DIRECTIONS

4.1. Conclusions

In Chapter 2 of this dissertation, we asked whether the nucleosome could offer absolute protection from BER-induced DSBs. To address this question, we incubated DNA glycosylases hNTH1 or hOGG1 (with or without AP-endonuclease) with nucleosomes engineered with two closely spaced lesions on opposite strands. We found that nucleosomes did offer substantial – but not absolute - protection from inadvertent DSB formation. Multiple factors influenced the degree of DSB formation, both physical and enzymatic. Lesions located only 2 bp from each other on opposite strands inhibited DSB formation in both naked and nucleosomal DNA, due to the proximity of the two binding sites. In nucleosomes, BER-induced DSB formation was limited by strand cleavage of the more occluded strand (Figure 1). Of interesting note was the observation that the rate of DSB formation by NTH1 on 601 nucleosomes exceeded what one would expect if the breaks on each strand were randomly distributed. This suggested cooperativity by the enzyme, wherein cleavage at the first lesion site somehow facilitates cleavage of the second. DSB rates also varied depending upon the type of lesion and enzyme tested. OGG1 induced essentially no DSBs on its own, while similar concentrations of NTH1 readily produced strand breaks. Related to this, both glycosylases were limited by their intrinsic lyase activity, which could be bypassed by APE1 to induce more DSBs. Nucleosomes remained intact during DSB formation and, depending upon DSB location and DNA sequence, could retain both pieces of cleaved DNA.
DNA. Taken together, these results suggest that in vivo, DSBs generated by BER would occur primarily in regions of the genome associated with elevated rates of nucleosome turnover or remodeling, as well as non-nucleosome-associated linker DNA regions.

In Chapter 3, we systematically examined the final step of BER on nucleosomes. Because LigIIIα-XRCC1 must completely encircle the DNA helix in order to complete ligation, histone-DNA contacts must at least be transiently disrupted to complete nick sealing. By moving a DNA strand-nick further towards the dyad, we reduced LigIIIα-XRCC1 activity by ~3-fold. This is concurrent with the nucleosome unwrapping model, wherein sites further within the nucleosome are exposed less frequently; consequently, repair enzymes that must wait for DNA unwrapping events are inhibited to a greater degree from processing damages located closer to the dyad. Thus, LigIIIα-XRCC1 activity must be limited by nucleosome unwrapping. We also found that ligation efficiency is inversely correlated with nucleosome thermostability, and that addition of extra-nucleosomal “linker” DNA did not enhance ligation (as had been the case with DNA ligase I in a previous study). Experiments with restriction endonucleases as well as gel shift experiments revealed that LigIIIα-XRCC1 binding does not irreversibly disrupt the nucleosome. This was in contrast to the conclusion made in (Odell et al., 2011), where disruption seemed to be the best explanation for 1) enhancement of pol β activity on nucleosomes by LigIIIα-XRCC1, and 2) the disappearance of an intact nucleosome band in native gels by increasing LigIIIα-XRCC1 concentrations. Instead, we demonstrate that the super-shifted complex that forms upon enzyme addition is not only
reversible, but is a ternary complex formed between LigIIIα-XRCC1 and the nucleosome in which ligation is taking place.

4.2. Future Directions

4.2.1. Could the nucleosome allow for complete clustered lesion repair?

Our publication on clustered lesion BER has generated some interest within the DNA methylation field. Multiple 5-methylcytosines located within a CpG site are staggered by one base pair on opposing strands. Thus, either damage to both 5-methylcytosines, or the demethylation process itself could make CpG sites a potential target for DSB formation via near-simultaneous BER. While the close proximity of the stagger would inhibit DSB formation (Chapter 2 and previous studies), could repair/demethylation go to completion on such a substrate? A recent *in vitro* study by Weber et al. processed an oligomer containing a 5-carboxylcystine CpG substrate (a demethylation intermediate) with thymine DNA glycosylase (TDG) in conjunction with purified BER enzymes in an attempt to complete the entire demethylation process. As one might expect, combining substrate with only TDG and APE1 resulted in some DSB formation. However, TDG combined with APE1, pol β, and LigIIIα-XRCC1 not only repaired a portion of the substrate, but failed to generate even minimal (<1%) DSBs in the process, indicating sequential lesion processing (Weber et al., 2016). Could nucleosome DSB suppression similarly allow for complete repair of clustered lesions? We have only looked at the first two steps in BER-induced DSB formation in
nucleosomes, but if all four enzymes were present simultaneously, DSBs might be prevented through either blocking of the second site during repair of the first site, or through substrate sequestration and handoff. Adding the nucleosome into this mix could generate one site that is preferentially repaired over the other, biasing the reaction towards sequential repair instead of simultaneous repair.

4.2.2. What happens after a BER-induced DSB occurs?

*In vivo*, we know that BER-induced DSBs ("BER-DSBs") happen, and we can artificially increase these DSB rates by overexpressing glycosylases in mammalian cells and exposing them to IR (reviewed in (Sage & Harrison, 2011)). Conversely, we also know that siRNA of glycosylases reduces DSB formation as a consequence of IR. However, we do not know how they are repaired, nor do we know where in the genome they occur. In the event of a BER-DSB, how is that damage handed off to the DSBR machinery? Is there a preferred DSBR pathway? Alt-NHEJ employs several of the same players involved in SSBR and BER (notably PARP-1 and LigIIα-XRCC1), so might this pathway be biased over NHEJ or HR? PARP-1 might already be localized to the site by way of interactions with both LigIIα and XRCC1, or a nearby clustered SSB. As well, due to protein-protein interactions, DNA end cleanup enzymes such as PNK, APTX, and TDP-1 may already be in place to deal with any additional "dirty break" moieties within the clustered damage. A quick preliminary experiment might be to first mix some or all the enzymes from BER and alt-NHEJ together with a simple clustered damage to see if complete repair is feasible *in vitro*. Going back to the cell, we could look at recruitment at
BER-DSBs using immunofluorescence in cell lines overexpressing NTH1, OGG1 or APE1 and exposed to IR (a non-BER-DSB control might be to treat normal cells with higher levels of IR to induce similar levels of DSBs, or perhaps use a radiomimetic compound such as bleomycin). Co-localizations to look for would be overlaps between BER enzymes, γ-H2AX formation, and early-response proteins from each of the different DSBR pathways.

4.2.3. Where do BER-induced clustered damages occur within the genome?

What about the location of BER-induced DSBs? How protective is chromatin to this particular type of damage? Endogenous DSBs localize to transcription start sites and sub-telomeric regions in dividing cells (Seo et al., 2012), but maybe BER-DSBs tend to be generated elsewhere? In these same cells, we could examine where BER-DSBs were being created relative to control cells generating mostly non-BER-DSBs. ChIP-seq or ChIP-on-ChIP could be performed using an anti-γ-H2AX antibody to determine where the breaks are occurring. However, γ-H2AX immunofluorescence has notable disadvantages: 1) γ-H2AX is often not recruited directly to the DSB, sometimes γ-H2AX is actually located >10,000 bp from the damage; and 2) heterochromatin is refractory to γ-H2AX labeling. A new technique known as BLESS (Breaks Labeling, Enrichment on Streptavidin and next-generation Sequencing), which involves direct, in situ labeling of DSBs followed by next gen sequencing to identify DSB locations with far higher resolution (Crosetto et al., 2013). As well, because protein is removed during the BLESS
procedure, presumably heterochromatin would not provide a barrier to labeling using this method.

4.2.4. Could single molecule studies show us how LigIIIα finds nicks in DNA?

Out studies with LigIIIα have led us to conclude that the enzyme is dependent upon nucleosome unwrapping to seal nicks within nucleosomes, but exactly how LigIIIα searches for nicks along DNA and nucleosomes is still unknown. As discussed in Chapter 3, most DNA-binding proteins begin their search for a substrate in a three-dimensional, diffusion-driven manner. Once a protein collides with DNA, however, the search switches to a more efficient, one-dimensional search along the DNA surface (von Hippel & Berg, 1989). This search can take place as tracking or “sliding” along the DNA helix, or can involve microscopic “hopping” along the DNA, wherein the protein follows one face of the DNA and jumps over the grooves of the DNA.

We currently do not know in what manner LigIIIα travels along DNA in search of nicks; nor do we know how it acts when it encounters a nucleosome: Does the enzyme bypass the nucleosome? Is it halted? Does it attempt to travel through the nucleosome? A single-molecule approach may be able to answer some of our questions. Recent studies by the Wallace group used single molecule fluorescence to visually observe the search mechanism of glycosylases on both undamaged and damaged DNA. Stringing up DNA strands between silica beads, they generated a “tight rope” of DNA stained with YOYO-1. They then interrogated these DNA tight ropes by adding hexahistidine-tagged
glycosylases to anti-hexahistidine antibody conjugated quantum dots. These quantum dots allow for visualization and measurement of single glycosylase molecules as they travel along the lengths of DNA. From this experimental setup, they made several key observations about the search mechanism of glycosylases (Dunn et al., 2011) (Nelson et al., 2014). Very briefly, the oxidative glycosylases tested tracked around the DNA helix at the rate of diffusion. However, these enzymes sometimes moved at a slower rate, possibly in a damage seeking mode. This slow diffusion mode was mediated by a specific (and usually aromatic) amino acid, evidence that this “wedge” residue was key in lesion recognition.

Our lab has recently been coordinating with Dr. Andrea Lee of the Wallace group in conjunction with Dr. Shane Nelson of the Warshaw group with the goal of incorporating FLAG-tagged nucleosomes into the glycosylase single molecule studies. However, this system could also be adapted for the study of LigIII. In this case, we would likely begin by engineering *E. coli* expressed LigIIIβ, which lacks the BRCT domain needed for XRCC1-interaction, with a C-terminal tag (FLAG, hexahistidine, etc.), as an N-terminal tag may interfere with the zinc finger structure. Adding nicks into our template would be achieved through treatment of our DNA substrate with commercially available single-strand nicking enzymes, or through incomplete ligation of concatemerized plasmids (which is typically an undesired artifact, but would not be in our case). In these experiments, we even could start with a DNA substrate lacking nucleosomes, and with or without nicks engineered, in order to inform our basic
knowledge of the LigIII search mechanism. Additional studies could compare the movement of LigIIIB to LigIIIA, or perhaps more interestingly, LigIIIA with and without XRCC1.

4.2.5. What happens when LigIIIA encounters a nucleosome?

Continuing from the above single molecule studies, we could incorporate nucleosomes into our DNA substrate to ask additional questions. For example, can LigIII bypass a nucleosome located on its DNA path? A recent report by the Finkelstein group utilized a similar single-molecule fluorescence method to examine the action of mismatch repair complexes Msh2–Msh3 and Msh2–Msh6 on DNA as well as DNA “roadblocks” such as nucleosomes (Brown et al., 2016). Their methods differ somewhat from the Wallace group, mainly in that the DNA substrate is strung between a fluid lipid bilayer and a chromium pedestal, lining up the DNA strands parallel to one another in arrays sometimes referred to as “DNA curtains”. Using this method, they generated λ-phage DNA curtains (containing MMR-specific damages) reconstituted onto triple-FLAG H2A tagged octamers in order to determine that Msh2–Msh3 can bypasses nucleosomes as well as DNA binding proteins while diffusing along a length of DNA. This study’s methods could be adapted to ask whether LigIII (or glycosylases, for that matter), could bypass a nucleosome. This particular study offers some potential protocol optimizations to the silica bead flow cell setup used previously to study oxidative glycosylases. For example, here they often label the DNA and nucleosomes after enzyme movement is recorded, in order to avoid difficulties generated by dyes or by bulky, quantum dot
labeled nucleosomes. As well, with this method we could observe whether our enzymes might jump between parallel DNA strands, as was observed with Msh2–Msh3.

Notably, DNA damages in the above study were not engineered specifically within nucleosomes, as the study was focused on nucleosome bypass and not nucleosome interrogation. As well, they used λ-phage DNA rather than DNA containing discrete nucleosome positioning sequences, such as 5S rDNA or the Widom 601 sequence. To ask how LigIII might interact with a nucleosome containing a nick, we would need to use a DNA substrate containing discrete nucleosome positioning sequences. Ideally we would modify these sequences slightly to incorporate single-strand nicking sites at predictable nucleosome positions in order to generate DNA tight ropes or curtains containing nucleosomes with specifically positioned, ligatable nicks.

4.2.6. Does LigIIIα exert torsion on DNA?

I briefly touched upon how some DNA binding proteins exert force on DNA in Section 1.1.6 (Nucleosome Energetics). Proteins that exert force upon DNA as they travel can generate torsion within the DNA. As an example, as RNA polymerase elongates a transcript, it generates positive superhelical density ahead of itself, while accumulating negative superhelical density behind it (Liu & Wang, 1987). Is it possible that LigIIIα exerts force on DNA in such a manner as well?
I experimented briefly with generating a circularized DNA construct, into which a nucleosome could be reconstituted with a nicked site located within. If LigIIIα generated torsion, the idea was that this construct would generate superhelical density as LigIIIα traveled its length. This torsion then would possibly inhibit LigIIIα in its “travel” into the nucleosome, resulting in reduced ligation efficiency, when compared to DNA with free ends. Unfortunately, I was unable to generate this construct due to technical difficulties. As well, this experimental setup does not have a positive control to compare to LigIIIα.

Instead, perhaps a molecular tweezer approach described in Section 1.1.6 would be a more direct (if more technically advanced) means of asking the same question. An example approach uses a 2006 DNA gyrase study as a template, although there are alternative ways to setup such an experiment (Gore et al., 2006). Here, DNA would be strung between a fixed coverslip and a magnetic bead. Within the DNA sequence, a nick (in this case, one that is unligatable to avoid LigIIIα acting upon it) is engineered next to a small, submicrometer, fluorescent bead. When enzyme is added, the nick allows for rotation of the bead in response to enzyme-generated torsion. This rotation can then be measured and the torsion exerted on the DNA calculated. Here, DNA gyrase serves as a positive control to ask whether or not LigIIIα is generating force upon DNA.

4.2.7. Does SNF2H/SMARCA5 influence nucleosome BER?

Very recently, it was discovered that kinase casein kinase 2 (CK2) phosphorylation of XRCC1 mediates interaction with the chromatin remodeler
SNF2H/SMARCA5 (Kubota et al., 2016). This study also demonstrated a requirement for SNF2H in H$_2$O$_2$-induced damage foci formation by XRCC1, as well as foci co-localization of SNF2H and XRCC1. SNF2H is an ISWI homolog that catalyzes sliding of the histone octamer in order to render DNA accessible (Aihara et al., 1998) (Fan et al., 2003). Might remodeling activity by SNF2H promote efficient nucleosome ligation by LigIIIα/XRCC1? We could add purified SNF2H to our in vitro nucleosome assays containing damaged nucleosomes with LigIIIα/XRCC1 (or any/all of the other BER enzymes) to look for stimulation of activity. In these experiments, we could also examine the effect XRCC1 phosphorylation state by generating mutants that either mimic or ablate phosphorylation at CK2 sites. If SNF2H was indeed stimulatory to nucleosome BER, we would ask whether this stimulation dependent upon the presence of XRCC1 and/or its phosphorylation state. It would be extremely interesting if SNF2H stimulation occurred and was specific to XRCC1 phosphorylation, as this would be the first chromatin remodeler demonstrated to specifically enhance BER on a nucleosome substrate.
(A) To access a damaged base within a nucleosome, a glycosylase must first bind to the lesion via the minor groove, and then flip the base out of the DNA helix through the major groove. “Outward” facing lesions are thus easier to process than “inward” facing lesions buried towards the histone octamer. (B) When two lesions on different strands are staggered by only 2 bp, the proximity of the two binding sites suppress DSB formation is both naked and nucleosomal DNA. (C) If one lesion is optimally rotationally oriented on the nucleosome, but the other is occluded, DSB formation is limited by cleavage of the slower strand. (D) Maximal DSB formation on nucleosomes occurs when both lesions are both optimally oriented away from the histone octamer and optimally spaced (~4-6 bp).
COMPREHENSIVE BIBLIOGRAPHY


and urine. *Proceedings of the National Academy of Sciences, USA, 87*(12), 4533-4537.


218


223


233


Williams, S. P., Athey, B. D., Muglia, L. J., Schappe, R. S., Gough, A. H., & Langmore, J. P. (1986). Chromatin fibers are left-handed double helices with diameter and
mass per unit length that depend on linker length. *Biophysical Journal, 49*(1), 233-248.


APPENDIX A: APLF DOES NOT IMPACT LIGIIIα-XRCC1 ACTIVITY

Aprataxin and PNK-like factor (APLF, also referred to as PALF or Xip1) is named for the shared forkhead-associated domain present in PNK and APTX. This domain targets target casein kinase 2 (CK2) phosphorylated sequences such as those found on XRCC1 and XRCC4 (Iles et al., 2007). APLF possesses endonuclease and 3’ to 5’ exonuclease activities, has been demonstrated to facilitate NHEJ in human cell culture (Li et al., 2011) and has been implicated in PARP-1 signaling and SSBR via its tandem zinc finger domain (Rulten et al., 2008). However, for our studies, we are particularly interested in APLF’s reported histone chaperone activity. APLF has the ability to bind histone H3/H4 tetramer via a C-terminal acidic domain, allowing APLF to dismantle tetrasomes (Mehrotra et al., 2011). Could APLF enhance BER on nucleosomes, either through its chromatin disassembly activity or its interaction with XRCC1?

We incubated 1 nM -46 nick nucs, 5 nM LigIIIα-XRCC1, and 5 or 100 nM APLF for 30 minutes at 37°C, removing aliquots at various time point (See Chapter 3: Experimental Procedures for detailed process of substrate preparation, nucleosome reconstitution, protein purification, and standard ligation assay procedure). APLF was expressed and purified from E. coli. LigIIIα-XRCC1 and APLF were diluted and pre-incubated together on for 5-10 immediately prior to the start of the assay. We observed no difference in ligation in either naked DNA or nucleosomes (Figure A-1).
Top: Assay time course. Ligation products migrate as 183 nt bands. The minor band denoted by an asterisk corresponds to unligated DNA located 3’ to the nick, and does not interfere with the quantification of ligation extent. Bottom: Nucleosome ligation calculated as a fraction of naked DNA ligation. We found no significant difference in ligation at when either 5 nM or 100 nM APLF was added, in either nucleosome or naked DNA (naked DNA graph not shown).

Figure A-1. APLF does not impact LigIIIα-XRCC1 Activity