Challenges for Base Excision Repair Enzymes: Acquiring Access to Damaged DNA in Chromatin

by

Chuxuan Li

B.S. in Chemistry, Nanjing University, 2014

A dissertation submitted in partial fulfillment of the requirements for the Degree of Doctor of Philosophy in the Department of Chemistry at Brown University

Providence, Rhode Island

May 2020
This dissertation by Chuxuan Li is accepted in its present form by
the Department of Chemistry as satisfying the dissertation requirement
for the degree of Doctor of Philosophy.

Date ___________ ____________________________
Sarah Delaney, Director

Recommended to the Graduate Council

Date ___________ ____________________________
Christopher T. Seto, Reader

Date ___________ ____________________________
Arthur R. Salomon, Reader

Approved by the Graduate Council

Date ___________ ____________________________
Andrew G. Cambell
Dean of the Graduate School
Curriculum Vitae

Chuxuan Li

324 Brook Street, Providence, RI 02912, USA

EDUCATION

**Ph.D. in Chemistry**
Brown University, Providence, RI
Thesis Title: Challenges for base excision repair enzymes: Acquiring access to damaged DNA in chromatin
Advisor: Prof. Sarah Delaney

**B.Sc. in Chemistry**
Nanjing University, Nanjing, China

SKILLS AND TECHNIQUES

**Skills**
- HPLC, FPLC, LC-MS, PCR, Gel electrophoresis, DNA footprinting techniques, DNA synthesis & purification, Protein expression & purification, Design and execution of enzymatic assays

**Software**
- LATEX, PyMOL, KaleidaGraph, KinTex Explorer, Microsoft (PowerPoint, Word, Excel, OneNote)

RESEARCH EXPERIENCE

**Brown University, Providence, RI**

*Research Assistant*  
Advisor: Prof. Sarah Delaney

- Examined the activity of base excision repair enzymes in the context of chromatin using kinetic methods and chemical footprinting techniques in a nucleosome core particle (NCP) model system
- Developed and optimized a quantitative platform to globally assess DNA glycosylase activity in NCP
- Prepared canonical histone proteins and histone variant proteins to evaluate the influence of variant(s) incorporation on DNA glycosylase activity in NCP
- Trained and mentored several undergraduate and graduate students through development and execution of independent research projects in the laboratory
- Responsible for equipment maintenance and laboratory management

**Brown University, Providence, RI**

*Research Assistant*  
Advisor: Prof. Jason Sello

Dec 2014 – Dec 2015
• Explored the function of ClpP peptidases in *Streptomyces* bacteria and their small molecule inhibitors for antibacterial drug development

**Nanjing University, Nanjing, China**

Feb 2013 – Jun 2014

*Undergraduate Research Assistant*

Advisor: Prof. Xinghua Xia

• Studied the capture and stimulated release of circulating tumor cells on thermal-responsive polymer-coated gold substrate

**University of California Los Angeles, Los Angeles, CA**

Summer, 2013

*Summer Research Student*

Advisor: Prof. Hsian-rong Tseng

• Investigated the application of thermal-responsive microfiber-coated microfluidic chips in isolating prostate circulating tumor cells for whole exome sequencing

**PUBLICATIONS**


• Bilotti, K., Kennedy, E. E., Li, C. and Delaney, S., “Human OGG1 Activity in Nucleosomes is Facilitated by Transient Unwrapping of DNA and Is Influenced by the Local Histone Environment”, *DNA Repair* 59, 1-8 (2017)

**CONFERENCES**

Gordon Research Conference on Nucleosides, Nucleotides and Oligonucleotides, Newport, RI  
Jun 2019

Gordon Research Conference on Mammalian DNA Repair, Ventura, CA  
Feb 2019

256th American Chemical Society Fall National Meeting, Boston, MA  
Aug 2018

Gordon Research Conference on Mammalian DNA Repair, Ventura, CA  
Feb 2017

**INVITED PRESENTATIONS**

**DNA Repair and Mutagenesis Seminars, Boston, MA**  
Jan 2019

Talk: Li, C. and Delaney, S. Histone H2A Variants Enhance Global Uracil Excision in Nucleosomes.

**The Providence Area Aging Research Forum, Providence, RI**  
Oct 2018

Talk: Li, C. and Delaney, S. The Initiation of Base Excision Repair in Nucleosomes Facilitated by the Incorporation of Histone H2A Variants.
AWARDS AND HONORS

ACS Editors’ Choice Article 2019
for publication “Histone H2A Variants Enhance the Initiation of Base Excision Repair in Nucleosomes”.
*American Chemical Society*

Honorable Mention 2018
for the best oral presentation in the Division of Chemical Toxicology at 256th ACS Fall National Meeting.
*American Chemical Society*

Doctoral Research Travel Grant 2017–2018
for graduate research travel. Once per fiscal year.
*Graduate School, Brown University*

William R. Potter Graduate Student Conference Travel Grant 2016–2018
for conference and research travel. Once per fiscal year.
*Chemistry Department, Brown University*

Conference Travel Grant 2016–2018
for graduate conference travel. Once per fiscal year.
*Graduate School, Brown University*

National Scholarship 2013
*Nanjing University*

SERVICE AND LEADERSHIP

Committee Member 2016–2019
Graduate Student Leadership Committee, Department of Chemistry, Brown University

Volunteer 2017–2018
Graduate Student Recruiting, Department of Chemistry, Brown University

Committee Member 2016–2017
Diversity and Inclusion Action Plan Committee, Department of Chemistry, Brown University

TEACHING AND MENTORING EXPERIENCE

Teaching Assistant, Brown University
CHEM0090 Kitchen Chemistry Spring, 2018
CHEM2870 Departmental Colloquia Fall, 2016
CHEM0330 Inorganic Chemistry I Fall, 2014
Acknowledgements

To my advisor, Professor Sarah Delaney, for letting me join the group and do fascinating research when I almost lost faith in pursuing my graduate study. Thank you for all your kindness, patience, and support during my time at Brown.

To my labmates, particularly Erin and Paul, for the laughs and great times inside and outside the lab. I miss the coffee and lunch we had together. I also thank my former group members for being very generous with their time in training me when I was just starting out in the lab and knew little about DNA repair.

To my friends who supported me from near and far. Thank you for listening, chatting, and making delicious food together.

To my parents, for your life-long love and understanding. I love you.

To my husband, Binyang, for your love and encouragement. Like you said, “Within the infinite space and time of the Universe, it is the most beautiful coincidence to meet you and live with you on the same planet”. I feel the same way.
Contents

List of Tables xiii

List of Figures xiv

1 Introduction 1

1.1 Introduction to DNA ................................. 2
1.2 DNA damage and nucleobase lesions ..................... 3
1.3 Base excision repair .................................. 3
1.4 DNA glycosylases .................................... 4
1.5 Nucleosome core particle (NCP) is the basic repeating unit of DNA packaging .............................. 5
1.6 Factors influencing DNA accessibility in NCP to BER enzymes .... 7
   1.6.1 Translational and rotational positioning ............... 7
   1.6.2 DNA sequence ..................................... 11
   1.6.3 Transient DNA unwrapping .......................... 13
   1.6.4 Histone tails ..................................... 15
   1.6.5 Histone post-translational modifications ............. 17
   1.6.6 Histone variant incorporation ....................... 20
   1.6.7 Chromatin remodeling ............................... 24
2 Human OGG1 activity in nucleosomes is facilitated by transient unwrapping of DNA and is influenced by the local histone environment

2.1 Abstract

2.2 Introduction

2.3 Materials and Methods

2.3.1 Oligonucleotide synthesis and purification.

2.3.2 Ligation of 145 mer oligonucleotides.

2.3.3 Glycosylase expression and purification.

2.3.4 Histone proteins preparation and octamer assembly.

2.3.5 Nucleosome core particle (NCP) reconstitution.

2.3.6 Hydroxyl radical footprinting.

2.3.7 Glycosylase kinetics experiments.

2.3.8 Molecular modeling.

2.4 Results

2.4.1 Formation of homogenous NCPs.

2.4.2 Hydroxyl radical footprinting defines rotational position of lesions.

2.4.3 Glycosylase activity in NCPs with off-dyad lesions.

2.4.4 Glycosylase activity in NCPs with H2B-modified octamers.

2.5 Discussion

2.6 Conclusion

2.7 Supporting information
3 Histone H2A variants enhance the initiation of base excision repair in nucleosomes

3.1 Abstract

3.2 Introduction

3.3 Materials and Methods

3.3.1 Oligonucleotide synthesis and purification

3.3.2 Histone preparation, NCP reconstitution and heat-shifting assay

3.3.3 Global assessment of UDG and SMUG1 activity

3.3.4 Single turnover kinetics experiments

3.3.5 Hydroxyl radical footprinting

3.3.6 Micrococcal nuclease digestion

3.3.7 Analysis of differential accessibilities of macroH2A octasomes and hexasomes to UDG and SMUG1

3.3.8 Molecular modeling

3.4 Results

3.4.1 Characterization of NCPs with global C to U substitution

3.4.2 Enhanced excision of U from NCPs containing H2A variants

3.4.3 Kinetics of U excision by UDG and SMUG1 from H2A variant NCPs

3.4.4 MacroH2A NCPs are a mixture of octasomes and hexasomes

3.4.5 MacroH2A octasomes and hexasomes both exhibit increased accessibility to UDG and SMUG1

3.4.6 Mapping of U sites in H2A.Z and macroH2A NCP models

3.5 Discussion
4 Global Repair Profile of Human Alkyladenine DNA Glycosylase on Nucleosomes Reveals DNA Packaging Effects

4.1 Abstract
4.2 Introduction
4.3 Materials and Methods
  4.3.1 Global εA-containing 145 mer oligonucleotide synthesis and purification.
  4.3.2 Ligation strategy to synthesize undamaged 145 mer oligonucleotides complementary to εA-containing strands.
  4.3.3 Formation of global εA-containing duplex.
  4.3.4 Reconstitution of global εA nucleosome core particles (NCPs).
  4.3.5 Hydroxyl radical footprinting.
  4.3.6 DNase I footprinting of NCPs.
  4.3.7 AAG kinetics on global εA-containing DNA.
  4.3.8 Molecular modeling of NCP.
4.4 Results and Discussion
  4.4.1 AAG activity on global εA-containing duplex DNA.
  4.4.2 AAG activity on global εA-containing NCPs.
  4.4.3 Relationships between AAG activity and solution accessibility of εA lesions in NCPs.
4.5 Conclusion
4.6 Supporting information
4.7 References
5 Conclusions and Future Perspectives

5.1 Conclusions ...................................................... 180
5.2 Future Perspectives ............................................. 184
5.3 References ....................................................... 186
List of Tables

1.1 Human glycosylases and their main lesion substrate(s)\textsuperscript{a} .......................... 6

2.1 $k_{\text{obs}}$ values determined for duplex DNA and NCP substrates. .................. 67

3.1 Kinetic parameters of U excision at site 110 from duplex DNA and NCPs. ................................................................. 106

4.1 $k_{\text{obs}}$ values for $\varepsilon$A excision by AAG\textsuperscript{a} .............................................. 155

4.1 $k_{\text{obs}}$ values for $\varepsilon$A excision by AAG\textsuperscript{a} .............................................. 156

4.1 $k_{\text{obs}}$ values for $\varepsilon$A excision by AAG\textsuperscript{a} .............................................. 157

4.2 Solution accessibilities of adenine positions in NCPs. .............................. 160
List of Figures

1.1 The chemical structure of DNA. .................................................. 2
1.2 Examples of nucleobase lesions caused by oxidation, alkylation, deamination and hydrolysis. ................................................. 30
1.3 Short-patch sub-pathway of BER. .............................................. 31
1.4 DNA packaging and NCP structure. ........................................... 32
1.5 Transient unwrapping of NCP. ................................................. 33
1.6 NCP octasome and hexasome species. ...................................... 33
2.1 Representation of the NCP structure and lesions used in this study. . 61
2.2 Representative native PAGE showing the formation of NCP. ......... 62
2.3 Hydroxyl radical footprinting of DNA in an NCP. ....................... 64
2.4 A representative denaturing PAGE showing the excision of U by UDG under single-turnover condition. ................................. 65
2.5 Single-turnover kinetic results for OGG1 and UDG in NCP substrates. 66
2.6 Electrospray ionization mass spectra of acetylated H2B. ............. 69
2.7 Single-turnover kinetic results for OGG1 on 8oxoG OUT in NCPs containing modified H2B. ........................................... 69
2.8 Single-turnover kinetics for the removal of off-dyad U lesions by UDG in tailless H2B NCPs. ................................................. 70
2.9 Molecular model of NCP-bound OGG1 ........................................ 72
3.1 Confirmation of global U incorporation into Widom 601 DNA. .... 95
3.2 Representative native PAGE of reconstituted NCPs. .............. 96
3.3 Heat-shifting assay on canonical and H2A variant NCPs. ......... 96
3.4 Solution accessibility of nucleosomal DNA characterized by hydroxyl radical footprinting. ................................................. 98
3.5 Denaturing PAGE visualization of UDG activity on duplex DNA and NCPs containing global C to U substitution. .................... 99
3.6 Denaturing PAGE visualization of SMUG1 activity on duplex DNA and NCPs containing global C to U substitution. ............. 100
3.7 Quantification of U excision from NCPs containing global C to U substitution after 60 min-incubation with UDG or SMUG1. ........ 102
3.8 Single-turnover kinetics time courses of UDG and SMUG1 acting on duplex DNA and NCPs containing a single C to U substitution at site 110. .......................................................... 105
3.9 Denaturing PAGE visualization of hydroxyl radical footprinting of DNA in canonical and H2A variant NCPs. ......................... 108
3.10 Quantification of hydroxyl radical footprinting of DNA in canonical and H2A variant NCPs. ................................................... 109
3.11 Micrococcal nuclease digestion of canonical and macroH2A NCPs. . 109
3.12 Differential accessibilities of U at site 110 in canonical and macroH2A NCP species to UDG and SMUG1. ......................... 111
3.13 Model of a macroH2A octasome and hexasome with global C to U substitution. .......................................................... 113
3.14 Model of an H2A.Z NCP with global C to U substitution. ........ 114
3.15 Ligation scheme for LS_U_{110} and CS 145 mer. ........................................ 123

4.1 Representations of εA:T bp, NCP, and DNA sequence. .............................. 135
4.2 Confirmation of εA substitutions for A in Widom 601 I and J strands. 149
4.3 Single-turnover kinetic time courses of εA excision by AAG on the J
    strand. ........................................................................................................ 151
4.4 Single-turnover kinetic time courses of εA excision by AAG on the I
    strand. ........................................................................................................ 152
4.5 AAG retains full activity for 180 min. ......................................................... 153
4.6 Quantitation of εA excision over time in global DUP. ................................. 154
4.7 Formation of global εA-containing DUP and NCPs. ..................................... 158
4.8 Relationship between AAG activity on and solution accessibility of εA
    positions in the NCP. ..................................................................................... 158
4.9 Non-biased incorporation of εA-containing DNA into NCPs. ...................... 159
4.10 Confirmation of εA substitutions for A in Widom 601 I and J strands. 162
4.11 DNase I footprinting on 601I DUP and NCPs. ............................................ 164
4.12 DNase I footprinting on 601J DUP and NCPs. ............................................ 165
4.13 DNase I activity in NCPs. ........................................................................... 166
4.14 Proximity of histone tails to HIGH positions with low AAG activity. ....... 169
4.15 Ligation strategy for complementary 145 mer oligonucleotides lacking
    damage. ........................................................................................................... 173
Chapter 1

Introduction

1.1 Introduction to DNA

Deoxyribonucleic acid, or DNA, encodes hereditary information and is found in nearly all living cells. Since its first discovery in 1869 by Friedrich Miescher, interest in the study of DNA has never diminished. In the early 1950s, the double helical structure of DNA was deciphered pioneeringly by James Watson, Francis Crick and Rosalind Franklin (1). Each strand of the double helix is a polynucleotide chain, in which monomeric nucleotides are connected by phosphodiester bonds. The nucleotide monomer is made of three essential components: a nitrogenous base, a deoxyribose sugar, and a phosphate group attached to the sugar (Figure 1.1). There are four primary nitrogenous bases: adenine (A), guanine (G), cytosine (C), and thymine (T); it is the distinct pairing arising from hydrogen bonds between A and T, and G and C, that holds the two complementary strands together and results in the double helical pattern of DNA.

![Figure 1.1 The chemical structure of DNA. The double helical structure of DNA is maintained by hydrogen bonds formed between nucleobases on the two adjacent strands. A always pairs with T, and G always pairs with C.](image-url)


1.2 DNA damage and nucleobase lesions

DNA is constantly assaulted by damaging agents deriving from both exogenous sources, such as air pollution, ionizing radiation, and cytotoxic chemicals, and endogenous sources, such as reactive oxygen and nitrogen species from normal cellular metabolism and the inflammatory response (2). Major types of DNA damage include double- and single-strand breaks, inter- and intra-strand crosslinks, mismatched base pairs, and chemical modifications of nucleobases. If unrepaired, DNA damage can be cytotoxic and mutagenic, acting as the primary molecular basis for cancer and aging (3).

Chemically-modified nucleobases, which we refer to here as nucleobase lesions or simply lesions, can originate from a diversity of processes (4). Examples of lesions deriving from oxidation, alkylation, deamination, and spontaneous hydrolysis are shown in Figure 1.2. So far, ~100 nucleobase lesions have been identified in vitro, and ~15 of them have been discovered in cellular DNA (5). If allowed to persist, nucleobase lesions can cause mutations by being mispaired by a DNA polymerase during replication to permanently change the genetic code or be cytotoxic by resulting in polymerase stalling and ultimately cellular apoptosis. This thesis will discuss some prototypic lesions in next chapters: 8-oxo-7,8-dihydroguanine (8oxoG) caused by the oxidation of G, uracil (U) caused by the deamination of C, and 1,N\textsuperscript{6}-ethenoadenine (εA) caused by the alkylation of A.

1.3 Base excision repair

Given the high frequency of nucleobase lesion occurrence (more than 2×10\textsuperscript{4} lesions occur per cell per day (6)), the wide range of lesions formed (5), and the detrimental
consequences they can lead to, it is of essential importance that cells efficiently repair the damaged DNA to maintain genomic integrity and cell viability. Indeed, several mechanisms have evolved that repair the majority of nucleobase lesions, including base excision repair (BER), nucleotide excision repair (NER) and non-homologous end joining (NHEJ). The BER pathway is specifically responsible for repairing non-bulky and non-helix-distorting nucleobase lesions, and is the focus of this thesis.

BER is a coordinated process, during which several enzymes work to replace the lesion with an unmodified nucleobase (Figure 1.3). The repair event is initiated by a DNA glycosylase that recognizes a specific lesion, cleaves the glycosidic bond which attaches the lesion to the sugar-phosphate backbone, and generates an AP site. APE1 (apurinic/apyrimidinic endonuclease 1) then cleaves the DNA backbone 5’ to the AP site, generating 3’-OH and 5’-deoxyribose phosphate (5’-dRP) termini. In short-patch BER (Figure 1.3), polymerase β (pol β) follows APE1 by removing the 5’-dRP and inserting the correct nucleotide at the 3’-OH. In the long-patch sub-pathway, in contrast, several nucleotides are inserted, creating a DNA flap that is next cleaved by FEN1 (flap endonuclease 1). Finally, a DNA ligase seals the nick in DNA backbone to complete the repair event (7).

1.4 DNA glycosylases

Since glycosylases are the first enzymatic players in BER, extensive effort has been devoted to understanding their mechanism of action and cellular functions (8–12). Eleven glycosylases have been identified in human cells, and they are grouped into six superfamilies according to their structural features (13). Table 1.1 shows each of the eleven glycosylases and their substrate specificity. Depending on whether they
can catalyze strand cleavage after excising the nucleobase lesion, glycosylases are classified as either monofunctional or bifunctional (14). Monofunctional glycosylases, for instance uracil DNA glycosylase (UDG/UNG) and AAG, generate an AP site. Bifunctional glycosylases, for instance OGG1 and NEIL1, generate an AP site and subsequently cleave the DNA backbone using their AP lyase activity. Interestingly, despite targeting different lesions, glycosylases (with the exception of those from the ALK family) share the same “pinch, push, plug, pull” mechanism to recognize, orient, and excise the lesion (13). Specifically, glycosylases search for lesions by probing the base pair stability and flipping the lesion into their active site. An amino acid residue fills the vacant space left behind by the extruded lesion to stabilize the DNA-enzyme complex, allowing for subsequent glycosidic bond cleavage.

1.5 Nucleosome core particle (NCP) is the basic repeating unit of DNA packaging

A large body of work has been carried out to define glycosylase activity using duplex DNA as a model system, where the DNA substrate is free in solution in the absence of other cellular components. However, eukaryotic DNA when packaged as chromatin presents a more complex substrate context for repair (15). The basic repeating unit of chromatin is the nucleosome, which is made of a nucleosome core particle (NCP) and linker DNA that connects adjacent NCPs (Figure 1.4A). An NCP consists of 145-147 base pairs (bp) of duplex DNA wrapped \(~1.7\) times around an octameric histone protein core (16). The protein core is comprised of two copies each of the four canonical histone proteins: H2A, H2B, H3 and H4. Within an NCP, there exists a dyad axis of pseudosymmetry that runs through the central base pair of nucleosomal
Table 1.1. Human glycosylases and their main lesion substrate(s)\textsuperscript{a}

<table>
<thead>
<tr>
<th>Glycosylase</th>
<th>Main substrates\textsuperscript{b}</th>
</tr>
</thead>
<tbody>
<tr>
<td>UNG</td>
<td>U, 5fU</td>
</tr>
<tr>
<td>SMUG1</td>
<td>U, 5hmU, 5fU</td>
</tr>
<tr>
<td>TDG</td>
<td>T\textsuperscript{=}G, U, 5fC, 5fU, 5caC, 5hmC, 5hU, Tg, \epsilon C</td>
</tr>
<tr>
<td>MBD4</td>
<td>T\textsuperscript{=}G, U, 5fU, \epsilon C, 5mC</td>
</tr>
<tr>
<td>AAG</td>
<td>\epsilon A, Hx, 3mA, 7mG, 3mG</td>
</tr>
<tr>
<td>OGG1</td>
<td>8oxoG, FapyG, FapyA</td>
</tr>
<tr>
<td>MUTYH</td>
<td>\textsuperscript{A}A:8oxoG, \textsuperscript{A}G, \textsuperscript{A}:C, 2hA:G</td>
</tr>
<tr>
<td>NEIL1</td>
<td>Tg, DHU, DHT, 5hU, 5hC, Gh, Sp, FapyA, FapyG</td>
</tr>
<tr>
<td>NEIL2</td>
<td>DHU, DHT, 5hU, 5hC, Gh, Sp, FapyG</td>
</tr>
<tr>
<td>NEIL3</td>
<td>Gh, Sp, FapyA, FapyG</td>
</tr>
<tr>
<td>NTH1</td>
<td>Tg, DHU, 5hU, 5hC, FapyG</td>
</tr>
</tbody>
</table>

\textsuperscript{a}Abbreviations for glycosylases and lesions that are not mentioned in previous text: UNG (uracil DNA glycosylase), SMUG1 (single-strand selective monofunctional uracil DNA glycosylase), TDG (thymine DNA glycosylase), MBD4 (methyl-binding domain glycosylase 4), AAG (alkyladenine DNA glycosylase), OGG1 (8-oxo-7,8-dihydroguanine glycosylase), MUTYH (MutY homolog DNA glycosylase), NEIL1/2/3 (Endonuclease VIII-like glycosylase 1/2/3), NTH1 (Endonuclease III-like 1), 5fU (5-formyluracil), 5fc (5-formylcytosine), 5caC (5-carboxycytosine), 5hmC (5-hydroxymethylcytosine), \epsilon C (ethenocytosine), 5mC (5-methylcytosine), 2hA (2-hydroxyadenine), DHT (5,6-dihydrothymine), and 5hU (5-hydroxyuracil).

\textsuperscript{b}X:X represents a mispair where the underlined nucleobase is excised by the glycosylase.
DNA (Figure 1.4B).

Due to the association with the histone core, DNA in an NCP is structurally constrained with limited dynamics, which challenges regulatory proteins that must gain access to nucleosomal DNA in multiple cellular processes including transcription and DNA repair. Several intrinsic elements of NCP have been shown to influence the stability and/or dynamics of NCP and the accessibility of a target site in nucleosomal DNA, such as the geometric location of the site, DNA sequence, transient unwrapping of nucleosomal DNA, and the presence of histone tails. Other mechanisms are utilized by cells to actively modulate genomic DNA accessibility, including post-translational modification of histones, substituting canonical histones with histone variant proteins, and remodeling chromatin structure by chromatin remodelers. In the following sections, how these factors influence DNA accessibility to BER enzymes within the context of chromatin will be discussed, with an emphasis on glycosylases that initiate the repair process.

1.6 Factors influencing DNA accessibility in NCP to BER enzymes

1.6.1 Translational and rotational positioning

The location and geometry of a nucleobase in an NCP can be described in terms of its translational and rotational positions. Translational position refers to the displacement of a nucleobase from the dyad axis (Figure 1.4B). Rotational position refers to the orientation of a nucleobase relative to the histone core (Figure 1.4C). Nucleobases facing outward and away from the histone core are highly accessible to solution, while
those facing inward toward the histone core are structurally sequestered and solution inaccessible (Figure 1.4D).

Significant effort has been devoted to investigating the influence of rotational positioning on the initiation of BER in NCPs. A general observation is that glycosylase activity correlates with rotational positioning (or solution accessibility) of a lesion. Specifically, it has been reported that U and oxidative lesions including 8oxoG, Tg, and Gh, can be excised from NCPs with efficiencies comparable to duplex DNA when the lesion is orientated away from the histone core (17–21). Conversely, minimal levels of lesion excision are detected at inward-facing sites. Notably, there are also exceptions where rotational positioning fails to dictate glycosylase activity, indicating additional regulatory elements exerting influence on lesion accessibility and repair, such as translational positioning (20) and histone tails (21).

The majority of these previous studies used NCPs containing one lesion at a defined location. In order to examine glycosylase activity on a lesion in a different geometric location another NCP had to be prepared individually. To gain an overall picture of the repair profile, strategies to globally incorporate a lesion into NCPs have been developed. Using PCR and a nucleotide pool containing the standard dNTPs and dUTP, the polymerase can incorporate T or U opposite a templating A during elongation. Since the dUTP concentration is much lower than that of dTTP, the DNA contains no more than one U:A bp per duplex. Using this strategy, collectively each T is replaced by U, which enables the NCP ensemble to possess U lesions at a variety of translational and rotational positions. Based on this platform, a correlation between rotational positioning and UDG activity in an NCP was observed, with the dyad region as a notable exception (22, 23). In other work, however, it was reported that rotational positioning alone is not predictive of lesion excision (24). In particular, some occluded sites exhibit high level of U excision by UNG, while
some solution-accessible sites exhibit limited amount of lesion excision. Further in-
spection revealed that the differential UNG activity at these sites is a comprehensive
consequence contributed by several elements, including the width of a minor groove,
presence of histone side chains, and dynamic flexibility of the lesion site (24), high-
lighting the intrinsic complexity of NCP as substrates for BER enzymes.

Another strategy for generating NCP with global incorporation of a lesion is
accomplished by chemical synthesis (25). In this case, a mixture of phosphoramidite
building blocks of lesion and the natural nucleobase is used, with the molar ratio
determined by a Poisson distribution to maximize the number of DNA strands with
one lesion. Compared to the PCR strategy, this method expands the scope of lesions
that can be examined, since the kinetics of incorporating some lesions during PCR by
a polymerase are too slow to effectively compete with the natural dNTP. Using this
chemical synthesis approach, it has been demonstrated that OGG1, UDG, and AAG
are capable of removing the lesion from 8oxoG:C, U:G, and εA:T pairs, respectively,
on a global level (25–27), emphasizing that DNA packaging is not a complete barrier
to BER. Detailed results of UDG and AAG repair will be discussed in later chapters.
Notably, experiments characterizing the global repair profile of glycosylases reveal
the architectural nuances of NCPs, for instance, transient unwrapping of nucleosomal
DNA and the interference of histone tails (27).

Translational positioning can also be a key factor in determining glycosylase
activity. Studies using NCPs containing site-specific or globally-incorporated lesions
both demonstrated a dramatic suppression of glycosylase activity in the dyad region
(20, 22, 25–28). Notably, unlike restriction enzymes that exhibit a $\sim10^4$-fold decrease
in activity as their recognition sites are moved from DNA ends to the dyad axis (29),
it appears that there is a boundary defining the region around the dyad axis where
glycosylase activity becomes significantly inhibited. It has been shown in vitro that
8oxoG, U, and εA are generally poorly repaired regardless of rotational positioning in the region ~20 bp region centered on the dyad axis (25–27). Importantly, these observations are in good agreement with data from in vivo studies where accumulation of oxidative and alkylation damage and high mutation frequency around the dyad axis are detected (30, 31), highlighting the imbalance between damage and repair that leads to mutations and genomic instability. One explanation for the diminished lesion accessibility in the dyad region is that DNA within ~30 bp centered on the dyad axis adopts a distinct helical periodicity that differs from that in other regions of the nucleosome (32, 33).

Interestingly, our laboratory has observed that U and εA at a specific site, located 2 bp away from the dyad axis and facing out toward solution, can be excised from NCP by UDG and AAG, respectively, as readily as from duplex DNA (34). Nevertheless, among all glycosylase/lesion pairs examined in the same work, including UDG/U, UDG/5hU, AAG/εA, AAG/Hx, OGG1/8oxoG, Fpg (formamidopyrimidine DNA glycosylase)/8oxoG, Nth (E.coli homolog of human NTH1)/5hU and Fpg/5hU, UDG/U and AAG/εA are the only two pairs that exhibit efficient repair at this particular site. Taking advantage of molecular modeling, it was suggested that the size of glycosylase, bending angle of DNA backbone, steric clash between glycosylase and the histone core, lesion type, and unique mechanism of glycosylases for processing lesion can all contribute to the differential levels of repair (34).

In regions that are over ~10 bp away from the dyad axis, glycosylase activity becomes more predictable based on rotational positioning. However, in the region ~20 bp from the DNA ends, it has been shown that inward-facing 8oxoG and Tg are more efficiently removed by OGG1 and NTH1, respectively, than those located closer to the dyad axis (21, 28). A generally-accepted view is that the transient dissociation of nucleosomal DNA from the histone core, starting from the DNA ends
(or DNA entry/exit region), confers higher dynamics and structural flexibility to nucleobases. This transient detachment of DNA is thought to be responsible for glycosylases obtaining access to lesions that would otherwise be structurally occluded. Of interest, recent work in our laboratory investigated the AAG repair profile using NCP with global εA lesions and described an overall increase in εA excision in a 20 bp region only at one DNA end (27), where the transient unwrapping is known to be thermodynamically favored. Details of this work will be presented in §4.

1.6.2 DNA sequence

In contrast to DNA-binding proteins that might target a specific sequence, such as restriction enzymes, the histone protein core has evolved to minimize sequence dependence. Nevertheless, the histone core does not bind to all DNA sequences in the same manner. Indeed, the binding affinities for the histone core can vary by greater than 1,000-fold based on DNA sequence effects (35). Importantly, this sequence selectivity is essential for defining nucleosome positioning and occupancy throughout the genome and has a profound influence on biological outcomes (36).

One of the major sequence features that determines NCP formation is the periodic placement of dinucleotides, TA/TT/AA, which are more bendable and stretchable to wrap around the histone octamer (37). A study of nucleosome positioning in genomic DNA of S.cerevisiae showed that TA/TT/AA dinucleotides occur every ~10 bp, and there exists a strong preference for these dinucleotides to reside in the minor groove facing towards the histone core (38). Similar observations were made when mapping the positioning of chromatin isolated from chicken erythrocytes (33).

Another sequence determinant in nucleosome formation are homopolymeric deoxyadenosine nucleotides on one strand of duplex DNA, referred to as poly(dA:dT)
tracts (39). They disfavor interaction with the histone core due to their intrinsic rigidity, and contribute to significant nucleosome depletion at these sequences in vivo (37). Of note, enrichment of poly(dA:dT) tracts is general in eukaryotic genomes and they serve as a functional element in the promoter region by preventing nucleosome formation and/or recruiting transcription proteins (40). Taken together, these studies highlight the importance of DNA sequence in forming NCPs, regulating histone-DNA interactions, and modulating nucleosomal DNA accessibility to cellular proteins.

As mentioned above, while most DNA sequences have some affinity for the histone core, some 145-147 bp sequences have been identified as having a stronger affinity. These DNA, termed nucleosome positioning sequences, bind to the histone core in a single and reproducible manner. Thanks to this feature, strong positioning sequences have been used widely to reconstitute NCPs in vitro, as they ensure a homogeneous population of NCP where the translational and rotational positions of nucleobases are predictable. Two such positioning sequences are the Widom 601 sequence and the 5s rDNA sequence.

The 5s sequence is a naturally-occurring sequence derived from the 5s ribosomal RNA gene of the sea urchin _L.variegatus_, and is highly conserved among organisms (41). Despite the fact that it can bind to the histone octamer with multiple translational positions in vitro (42), incubating the NCPs at elevated temperature induces re-positioning of the nucleosomal DNA with respect to the histone core, giving rise to a uniform and thermodynamically-stable translational position (43). The 5s sequence has served as a model sequence used in a number of biochemical investigations, such as scrutinizing NCP architecture by DNase and hydroxyl radical footprinting (32, 44), characterizing the interplay between nucleosome, transcription factor and linker histone H1 (45), and defining glycosylase activities in the context of chromatin (28, 46, 47).
When considering BER in NCPs many other studies have been performed using the Widom 601 sequence (20, 21, 24, 34, 48). This DNA was originally identified from a library of chemically-synthesized DNA fragments as a sequence that has a very high affinity for the histone octamer (49). Although the 601 sequence is not found in biology, NCPs reconstituted using the 601 sequence are homogenous and structurally stable under physiological salt concentrations and temperatures, providing a well-defined model system for \textit{in vitro} studies. Compared with the 5s sequence, NCP bearing the 601 sequence are less dynamic in solution and thermodynamically more stable (50, 51). Additionally, the reported crystal and cryo-EM structures of 601 NCPs provide detailed structural information, which can inform biochemical studies (52, 53).

\subsection*{1.6.3 Transient DNA unwrapping}

As mentioned in §1.6.1, transient unwrapping occurs prevalently at DNA entry/exit regions (Figure 1.5). Using restriction enzymes as probes, earlier studies from Widom and colleagues showed that transient unwrapping contributes to the accessibility of target sites that would otherwise be sterically occluded (54), proposing the probability of DNA-processing enzymes competing with histone proteins for nucleosomal DNA binding. They also determined that exposure rates of target sites is sequence dependent and diminish substantially as a function of distance away from the DNA ends (29). Of note, this decrease in site accessibility is manifested with glycosylase studies as well, in which inward-facing lesions are observed to be more efficiently removed when positioned close to the DNA ends than at the dyad axis (20, 21, 28), suggesting transient unwrapping as a potential regulatory factor of BER in the context of packaged DNA.
Interestingly, transient unwrapping can occur asymmetrically with one end dissociating from the histone core preferentially. The asymmetry of nucleosomal DNA unwrapping has been demonstrated using single-molecule fluorescence-force spectroscopy, time-resolved small-angle X-ray scattering and fluorescence resonance energy transfer (FRET), cryo-EM, and computational simulation (55–58). These studies show that the source of asymmetric unwrapping derives from the unequal free energy of histone-DNA interaction at the two ends, which is intrinsically determined by local DNA flexibility (or DNA sequence). Furthermore, dissociation of nucleosomal DNA at one end stabilizes histone-DNA interaction at the other end. Importantly, asymmetric unwrapping occurs not only under experimental conditions such as applied external force or with salt-induced disassembly (55, 56), but also under physiological conditions (57). Intriguingly, the impact of asymmetric unwrapping can propagate through the NCP by inducing conformational changes within the octamer core, particularly in H3, and rearrangement of H2A-H2B interaction with nucleosomal DNA (57). It is plausible that asymmetric unwrapping may be exploited by cells to regulate NCP interacting with other cellular components, such as ATP-dependent chromatin remodelers, histone chaperones, and transcription and DNA repair machinery (57, 59).

The influence of asymmetric unwrapping has been observed on BER enzymes in our recent work investigating AAG repair on NCP (27). Using an NCP ensemble that contains global εA lesions, the end that preferentially unwraps has been observed to have efficient lesion excision irrespective of rotational positioning. §4 will discuss more about this observation and architectural nuances of NCP revealed by the global AAG repair profile.

While it has been demonstrated that unwrapping can occur, in order to understand its potential biological impact, one must consider the timescale of this
dynamic motion. Results obtained from stopped-flow FRET experiments reported rate constants for spontaneous unwrapping and rewrapping of nucleosomal DNA to be $\sim 4 \text{ s}^{-1}$ and $\sim 20$-$90 \text{ s}^{-1}$, respectively (60). Other FRET studies determined the unwrapping and rewrapping rate constants to be $\sim 210 \text{ s}^{-1}$ and $370 \text{ s}^{-1}$ (61). The discrepancy between the reported rate constants is mostly likely due to the different salt concentrations used in experiments, since salt concentration can dramatically influence the electrostatic interaction between histones and DNA, and nucleosomal DNA dynamics. The reported rate constants indicate the lifetime of the partially unwrapped state of NCP ranges between $\sim 3$-$50 \text{ ms}$, which may enable the binding of DNA-processing proteins to NCP in biological events that occur on a comparable timescale (for a detailed review of nucleosome unwrapping and dynamics, see (59)).

As for BER enzymes, a study using the glycosylase NTH1 as a probe estimated that unwrapping-mediated exposure of an oxidative lesion $\sim 16 \text{ bp}$ from the DNA end occurs $\sim 7$-$8$ times per minute, which is likely too slow to account for the observed rates of NTH1 in vitro (47).

Importantly, it has been shown that the unwrapping can occur not only in single NCP, but also in systems of higher order compaction, such as nucleosome arrays. This dynamic motion of nucleosomal DNA in arrays allows for binding of a transcription repressor (62, 63). Moreover, binding of the repressor further induces the decompaction of nucleosome arrays, indicating the role of transient unwrapping in modulating chromatin dynamics and facilitating protein binding in vivo.

### 1.6.4 Histone tails

The core histones H2A, H2B, H3 and H4 contain two common structural domains: the central histone fold and the unstructured, lysine-rich tails (16). The N-terminal tails
of the four core histones and the C-terminal tail of H2A protrude from the NCP and are highly flexible in solution (Figure 1.4C) (64, 65). In addition to interacting with nucleosomal DNA within the same NCP, histone tails also associate with linker DNA (66) and the acidic patches of adjacent NCPs (16, 67). It has been shown that the N-terminal tails of H2A, H2B, H3, and H4 are all functionally important for forming the higher order chromatin structure (67–71). Deletion of histone tails has been shown to increase the accessibility of nucleosomal DNA and promote transcription factor binding. This binding is likely due to the enhanced nucleosome dynamics and reduced nucleosome stability in the absence of the tails (72–75). Importantly, proteolytic clipping of histone tails has been demonstrated in vivo and linked to gene expression, cell differentiation, aging, and disease development (74, 76, 77). Furthermore, histone tails are hotspots for post-translational modifications (PTMs), and they play essential roles in a multitude of cellular events including DNA repair (see review (78)), which will be discussed in §1.6.5.

Due to their distinct structural and functional features, histone N-terminal tails have piqued extensive interest in investigating their participation in BER (79). It has been reported that lack of histone N-terminal tails has no effect on the activity of UDG, APE1, and pol β in repairing U at a position near the dyad axis within an NCP (80). A similar observation has been made in our work, where removal of the H2B tail does not influence OGG1 excision of 8oxoG near the H2B tail that protrudes between the superhelices (more of this work will be discussed in §2) (21). In contrast, histone tails have been shown to interfere with ligase I (81). While ligation is severely inhibited at a nick site positioned close to the dyad axis in an NCP, introducing linker DNA or removing histone tails can both alleviate the inhibition of ligase activity. These findings suggest that the presence of linker DNA re-directs the interaction between histone tails and nucleosomal DNA (82, 83); otherwise the
histone tails would associate with ligase and block its access to the nick site (81). Interestingly, the opposing effect of histone tails has been observed for FEN1, which exhibits lower activity in a tailless NCP, likely due to the tails orienting the DNA flap to facilitate recognition and binding by FEN1 (84).

In addition to structurally modulating BER enzyme activity, histone tails can affect the lifetime of lesions within an NCP. It has been shown that rapid crosslinking between AP sites (and C4'-oxidized AP sites) and lysine residues in the H4 tail significantly shortens the lifetime of AP sites and generates strand breaks (85, 86). Furthermore, the reactivity of AP sites in forming protein–DNA crosslinks varies with the location of the lesions and their proximity to the histone N-terminal tails (87). Altogether, the fact that histone tails can both facilitate and inhibit the activity of BER enzymes underscores the complicated roles of histone tails in DNA damage repair and maintaining genomic stability.

1.6.5 Histone post-translational modifications

Post-translational modifications (PTMs) of proteins play essential roles in a multitude of biological processes (88–90). These modifications can function as marks for recruiting binding factors or by inducing protein conformational changes and altering their physical properties. An astonishing number of PTMs of histones have been identified including but not limited to acetylation, phosphorylation, methylation, ubiquitination, sumoylation, and ADP-ribosylation (91). The elusive pattern of histone PTMs, hypothesized as the histone code, has drawn tremendous interest due to its correlation with distinct chromatin states that regulate DNA accessibility.

Since the discovery of the first acetylated lysine in histones and its role in regulating gene transcription (92), histone acetylation has been the most intensively
investigated modification. This modification is installed by histone acetyltransferases (HATs), which catalyze the transfer of an acetyl group to the ε-amino group of lysine side chain using acetyl-CoA as a cofactor. Two mechanisms have been proposed by which histone acetylation modifies chromatin structure and function. Given that the interaction between histones and DNA is predominantly electrostatic, one mechanism emphasizes the fact that acetylation neutralizes the positive charge on lysine and thus confers a local, reduced interaction between histones and DNA. As a result of this weakened histone-DNA interaction, chromatin adopts a more flexible configuration or exhibits higher dynamics depending on the location of the acetylated lysine (for a detailed review of PTMs influencing nucleosome structure and dynamics, see reference (93)). For instance, acetylation of lysine 56 in histone H3 (H3K56ac), which resides near the DNA entry/exit region, has been shown to significantly facilitate the transient unwrapping of nucleosomal DNA to increase DNA accessibility and transcription factor binding (61, 94, 95). On the other hand, as the histone-DNA interaction at the dyad axis has been shown to be the strongest and critically important for NCP stability (96), acetylation of lysine near the dyad axis, such as H3K115ac and H3K122ac, destabilizes NCP to induce disassembly under external force (97) or by chromatin remodelers RSC (remodel the structure of chromatin) and SWI/SNF (switch/sucrose-non-fermenting) (98).

The other mechanism for histone acetylation altering chromatin structure and function refers to the acetylation acting as a mark to recruit bromodomain-containing proteins. One example is the recognition of K14ac in the N-terminal tail of H3 by the Rsc4 subunit of ATP-dependent chromatin remodeler RSC (99). Rsc4 bears tandem bromodomains targeting H3K14ac to assist the RSC complex in gene activation. Notably, the mechanisms of neutralizing the lysine positive charge and recruiting proteins to the acetylated lysine can work synergistically. It has been shown
that H3K14ac is essential in fission yeast for DNA damage checkpoint activation by directly regulating the compaction of chromatin and by recruiting RSC (100). Furthermore, hyperacetylation at H3K9 and H3K14 has been observed in yeast after UV irradiation accompanied by chromosomal DNA becoming more accessible (101). The level of histone hyperacetylation diminishes gradually as repair proceeds, indicating the essential role of histone acetylation in NER.

Despite in-depth studies of the effect of histone acetylation on NER and double strand break (DSB) repair (see review (102)), functions of histone acetylation in BER have not been well characterized. Our work has shown that substitution of canonical H2B in NCP with a chemically-acetylated counterpart leads to an enhancement in 8oxoG excision by OGG1 near the H2B tail (21). In contrast, depletion of H2B tail does not affect lesion excision. These observations indicate that it is the positive charges of lysine in H2B, rather than steric obstruction by the H2B tail, that predominantly influences OGG1 reactivity in NCP. §2 will have further discussion about these observations. In other work, acetylation of H3K14 or H3K56 did not contribute to U excision by UDG, regardless of the translational or rotational position of the lesion within an NCP (103). However, the gap-filling activity of pol β is suppressed significantly in the presence of H3K56ac or H3K14ac. Though the mechanism by which histone acetylation modulates the reactivity of repair enzymes in NCPs is not fully understood, these results show that histone acetylation can both facilitate and inhibit the activity of BER enzymes.

Histone methylation at lysine residues is another common PTM that is accomplished by histone methyltransferases (104–106). Several lysines in H3 and H4 can be methylated, including H3K4, H3K9, H3K27, H3K36, and H4K20 (107). Methylation of H4K20 has been shown to signal for recruiting checkpoint protein Crb2 to the damage sites; mutation of H4K20 significantly increases cell vulnerability to
DNA damage induced by UV light and ionizing irradiation (108). Nevertheless, little is known about the potential role(s) of histone methylation in BER. Recent work showed that a mono-methylated H4K20 mimic lead to a shorter lifetime of AP lesions near the H4 tail within an NCP (109). Compared to canonical NCP that bears no PTMs, NCP containing the H4K20me mimic exhibits an accelerated β-elimination step after DNA-protein crosslink formation between the mimic and the AP site.

1.6.6 Histone variant incorporation

Histone variants are nonallelic isoforms of canonical histones. Compared to their canonical counterparts, histone variants exhibit varied degrees of divergence in amino acid sequence. The exchange of histone variants for canonical histones, which is accomplished by histone chaperones, can confer distinct structures, dynamics, and PTM patterns to NCPs, thereby impacting multiple cellular processes. While canonical histones are expressed and deposited during S phase in a replication-coupled manner, expression and incorporation of histone variants into chromatin is replication-independent, allowing for specialized spatiotemporal functions throughout the cell cycle (110). Extensive evidence has shown that histone variants play important roles in regulating transcription, epigenetics, DNA damage repair, and cell development, and that mis-regulation of and mutations in histone variants can lead to cancer and other diseases (reviewed in (111)).

In recent years, significant progress has been made in elucidating the role(s) of histone variants in the DNA damage response (DDR), especially DSB repair and NER. One of the most well-studied phenomenon is the genome-wide phosphorylation of H2A.X upon formation of DSB (112). Phosphorylation of H2A.X is an early event in DDR, which gives rise to remodeling of the chromatin landscape to facilitate
downstream repair processes. H2A.Z is another H2A variant that has been shown to be incorporated at DSB sites by p400 remodeling ATPase to promote acetylation of H4 and chromatin ubiquitylation (113). These chemical modifications result in a more open conformation of chromatin, which increase the physical accessibility of DNA to repair machinery. A similar mechanism is used by H2A.Z to facilitate efficient NER, during which H2A.Z occupancy promotes H3 acetylation for recruiting repair enzymes (114). Using chromatin immunoprecipitation (ChIP) a third H2A variant, macroH2A, has been shown to accumulate at DSB sites (115). Interestingly, whereas formaldehyde crosslinking treatment followed by ChIP analysis resulted in a strong signal for macroH2A, in the absence of the crosslinking macroH2A was not observed, indicating that macroH2A is not incorporated into nucleosomes during DDR. Instead, it is thought that macroH2A associates transiently with chromatin through the interaction of its distinctive macro domain with PARylated chromatin.

In addition to the versatile H2A variant family, H3 has two variants, H3.3 and centromeric H3 variant CENP-A, that are highly conserved throughout eukaryotes and are engaged in DDR (116, 117). Rapid recruitment of CENP-A to DSB has been observed and CENP-A accumulation is enhanced at the damage sites by active NHEJ, one of the two major pathways of DSB repair (118). On the other hand, H3.3 has been reported to be proteolytically cleaved at its N-terminal tail and the resulting truncated form is incorporated into NCPs during cellular senescence induced by oncogene activation or DNA damage (119). Taken together, these observations reveal the sophisticated, yet to be fully elucidated mechanisms by which histone variants signal DDR and senescence, re-organize chromatin architecture, modulate DNA accessibility, and assist in loading repair factors around the damage sites.

Despite the well-established links between histone variants and both DSB repair and NER, it remains largely unknown whether histone variants are players in
BER. Work from Menoni et al. investigated OGG1 excision of 8oxoG from H2A.Bbd (Barr body deficient)-containing NCP relative to canonical NCP (48). H2A.Bbd shares only 48% identity to canonical H2A (120). It has been discovered to be enriched in transcriptionally active regions of the genome and is largely excluded from the inactive X chromosome (121). Biophysical and biochemical characterizations of H2A.Bbd-containing NCP, including fluorescence recovery after photobleaching (FRAP), FRET, analytical ultracentrifugation, and micrococcal nuclease digestion, suggest that these NCP wrap 118-130 bp of DNA and that they adopt a more relaxed and less stable conformation than canonical NCP (122–124). These findings could tempt one to speculate that DNA in H2A.Bbd NCP would be more accessible to BER enzymes. However, an 8oxoG lesion positioned near the dyad axis and which exhibited intermediate solution accessibility, was inefficiently excised (less than 20%) in both canonical and H2A.Bbd NCPs (48). Intriguingly, in the presence of the SWI/SNF chromatin remodeler 8oxoG excision from canonical NCP is as efficient as from duplex DNA; in contrast, H2A.Bbd NCP are resistant to SWI/SNF-induced remodeling and minimal improvement in 8oxoG repair is observed.

Interestingly, macroH2A has been implicated in BER. Cells lacking macroH2A exhibit reduced viability upon methyl methanesulfonate (MMS) exposure, a methylating agent that creates lesions known to be repaired by BER (115). Our own biochemical studies have also shown that incorporating macroH2A into NCP can modulate BER by increasing U excision efficiency at most structurally occluded sites (26). Similar observations were made for H2A.Z NCP. Details of these findings will be presented and discussed in §3.

Previous crystallographic and biochemical studies suggest that macroH2A
NCPs have increased stability due to the altered L1-L1 interface between two macroH2A-H2B dimers, which is one region where macroH2A differs from H2A (125, 126). However, the observation of facilitated excision of U from macroH2A NCP highlights the impact of structural nuances of variant NCP and the capability of variant incorporation on modulating glycosylase reactivity in the context of packaged DNA. The stability of an NCP may not necessarily reflect the local structural microenvironments of lesions within chromatin. Of interest, in the same study U sites with the largest increase in UDG efficiency were found to be clustered around the docking domain residing at the C-terminus, the N-terminus, and the L1 loops of macroH2A and H2A.Z. These regions, together with the acidic patch, are places where variants in H2A family exhibits the highest degree of sequence divergence and distinctive structural features (127), emphasizing the idea that structural alterations in variant NCPs can be strategically controlled to regulate the repair process. More importantly, previous molecular dynamics simulations (128) and a recent study of H2A variants in *Arabidopsis thaliana* (129) suggest that the impact of the unique docking domain and L1 loops of H2A variants can alter not only the structure and dynamics of individual NCP, but also inter-nucleosomal interactions. Therefore, incorporating H2A variants may be profoundly influential on the accessibility of chromatin fibers to BER machinery.

Intriguingly, incorporation of CENP-A into chromatin has been linked to the activity of UDG. CENP-A co-localizes with UDG during G2 phase and inhibiting UDG activity results in a lack of detectable of CENP-A foci on *Xenopus* sperm chromatin (130, 131). Further studies of histone variants coupled with other cellular factors such as PTMs of the variants and chromatin remodelers, will provide a more comprehensive understanding of how histone variants can modulate BER efficiency and genomic stability.
1.6.7 Chromatin remodeling

In addition to histone PTMs and variant(s) incorporation, reorganizing the chromatin landscape by ATP-dependent chromatin remodeling complexes (or chromatin remodelers) is another major mechanism of modulating DNA accessibility. Depending on the different structural domains flanking a conserved ATPase domain, chromatin remodelers are grouped into four families, i.e. the SWI/SNF, CHD (chromodomain-helicase-DNA binding), ISWI (imitation switch) and INO80 (inositol requiring 80) family (132, 133). Using the energy of ATP hydrolysis, chromatin remodelers are capable of altering and disrupting DNA-histone contacts by driving nucleosomes to slide along DNA, or evicting/exchanging histones (134–136). They make DNA and histones available to regulatory proteins, and thus play pivotal roles in multiple cellular processes, including transcription, replication, and DNA damage repair (137–139).

There is increasing evidence implicating the participation of chromatin remodelers in BER. It has been shown in vitro that the activity of OGG1, APE1, and pol β are stimulated by SWI/SNF complex in repairing an 8oxoG lesion at a position close to the dyad axis in the NCP (48). A similar observation has been made that the barrier imposed by histones is greatly alleviated by ISW1 and ISW2 complex for pol β synthesis (140). Furthermore, an overall increase in U excision by UDG across NCP, especially at inward-facing sites, is achieved by chromatin remodeler RSC working synergistically with histone chaperone FACT (facilitates chromatin transcription) (23). In other work, RSC has been found to enhance OGG1 excision of 8oxoG that is positioned either in the linker DNA or within NCP (141). Interestingly in this study, although histone chaperone NAP1 (nucleosome assembly protein 1) can enhance the accessibility of 8oxoG in linker DNA by removing histone H1 from nucleosomal DNA,
itself is not sufficient for an efficient excision of 8oxoG within NCP without the remodeling activity of RSC. The findings of chromatin remodelers cooperating with histone chaperones highlight the sophisticated mechanisms utilized by cells to regulate BER.

The participation of chromatin remodelers in modulating BER has also been indicated by \textit{in vivo} studies, although detailed mechanisms need to be further elucidated. One member of the SWI/SNF family, CSB (cockayne syndrome group B) (142), has been shown to function in the repair of 8oxoG lesions (143, 144) as revealed by 8oxoG accumulation in CSB null and helicase mutant cell lines relative to wild-type cells after damage induced by \(\gamma\)-radiation (143). Furthermore, CSB co-localizes with OGG1 after \(\gamma\)-radiation, and its knockdown and mutations result in a reduced level of OGG1 expression and 8oxoG repair (144). The role of CSB influencing NEIL1 behavior has also been assessed (145). Specifically, it has been found that mice lacking CSB exhibit higher levels of FapyG and Fapy A, which are mutagenic lesions targeted by NEIL1. Additionally, CSB and NEIL1 co-immunoprecipitate in HeLa cells and CSB stimulates NEIL1 activity \textit{in vitro}. Similar observations have been made for NEIL2, particularly in a DNA bubble structure during transcription (146). Altogether, these findings suggest that CSB functions in response to oxidative stress and may cooperate with NEIL2 in transcription-coupled BER. Indeed, it has been shown that NEIL2 activity can be linked to transcription, as mice lacking NEIL2 accumulate oxidative damage in transcribed genes and are susceptible to inflammatory agents (147). Moreover, NEIL2 associates with RNA polymerase II (RNP II) and hnRNP-U (heterogeneous nuclear ribonucleoprotein-U), both \textit{in vitro} and in cells. NEIL2 immunocomplexes from cell extracts preferentially repair 5hU in the transcribed strand; when NEIL2 is depleted, more nucleobase lesions are accumulated in active than in silent genes (148).
Other studies have shown that depletion of subunits of RSC and INO80 complexes increases the sensitivity of yeast cells to MMS (149–151). Specifically, for RSC it has been shown that yeast lacking STH1, a critical ATPase subunit of the RSC complex, exhibit genome-wide BER deficiency in removing methylated nucleobases resulted from MMS treatment. Unlike the influence exerted by CSB on OGG1, depletion of STH1 does not change the expression of BER genes; instead, chromatin shows reduced accessibility to MNase (Microccocal nuclease) digestion in the absence of STH1 (151).

Working together with chromatin remodelers, histone chaperone proteins play important and diverse roles in regulating nucleosome assembly and DNA accessibility (152), one of which is the FACT complex. FACT is a highly conserved histone chaperone complex among eukaryotes (153), and functions centrally in nucleosome assembly and chromatin barrier modulation (154). In addition to its well-appreciated role in facilitating transcription, FACT has been implicated in transcription-coupled NER and DSB repair (155, 156). As for its involvement in BER, an interesting switch of FACT from binding with transcription-associated factors to binding with DNA repair enzymes, including DNA ligase III, XRCC1 (X-ray repair cross complementing 1) and PARP1 (poly(ADP-ribose) polymerase 1), after occurrence of H$_2$O$_2$-induced oxidative damage has been reported (23). Further in vitro experiments showed that FACT, when acting in concert with RSC, was capable of promoting UDG excision of U lesions in the NCP, particularly at sterically occluded positions. This global enhancement of U excision is attributed to FACT assisting RSC remodeling the NCP to a more relaxed structure, termed remosome (157), that has weaker DNA-histone interactions and exhibits increased DNA accessibility.

Interestingly, FACT is capable of facilitating RNP II-driven transcription by generating a nucleosome hexasome species (158). This subnucleosomal species
contains the \((\text{H3-H4})_2\) tetramer and only one H2A-H2B dimer, with the hexameric histone core associating tightly with \(\sim 110\) bp of nucleosomal DNA 1.6 (159, 160). There is a growing body of evidence showing that the hexasome is an important intermediate of nucleosome remodeling during transcription, replication, and repair (161). Considering the fact that NEIL2 associates with RNP II and preferentially repairs lesions in actively transcribed genes (148), it is intriguing to speculate that the hexasome formed during transcription modulates the chromatin barrier not only for transcription machinery but also for BER enzymes. Indeed, we have recently observed that the formation of nucleosome hexasome species can facilitate U excision in NCP by dramatically elevating DNA accessibility to glycosylases (26). Details of this work will be discussed in §3.

1.6.8 Higher order packaging

While a majority of \textit{in vitro} studies examined BER enzymes using NCP as a simplified model, attempts have also been made to reconstitute the BER process with more cellular components such as linker DNA and histone H1, mimicking biological environments. It has been reported that glycosylases can excise lesions in linker DNA with efficiencies comparable to free duplex DNA in both mono- and di-nucleosome substrates (22, 141). However, in the presence of histone H1, reduced rates of lesion excision in linker DNA and the dyad region have been detected. This reduction is because H1 associates with NCPs predominately at the dyad axis and in the linker DNA region, conferring a more stabilized and compact conformation to the NCP (53). To gain a more in-depth molecular understanding of how DNA packaging regulates BER, a 12-mer nucleosome array model with precisely positioned U has been designed, mimicking the higher order chromatin folding (162). Using this array, assessment of
U removal was performed at a low and high Mg$^{2+}$ concentration, representing an extended beads-on-a-string state or highly compacted state, respectively. Excision of a U lesion, which faces out and is in the core region of an NCP, was observed to be fastest in the mono-nucleosome, intermediate in the extended state, and the slowest in the compacted state of the nucleosome array. In contrast, when U was positioned in a linker DNA region, an opposite trend was observed among the three substrates, revealing the elusive effect that inter-nucleosome interactions exert on BER within higher order chromatin structures.

Furthermore, genome-wide mapping of methylation damage, repair, and mutagenesis in yeast at single-nucleotide resolution has been carried out (30). Consistent with in vitro studies, translational and rotational positioning significantly influenced BER in NCPs. Importantly, MMS-induced mutations at adenines were remarkably enriched in the non-transcribed strand, particularly in BER-deficient strains.

1.7 Concluding remarks and gap in knowledge

DNA is vulnerable to a plethora of chemical modifications that give rise to a wide variety of damaged nucleobases. The BER pathway is charged with the task of searching the genome for non-bulky nucleobase lesions and to repair them to maintain genomic stability and cell viability. A majority of in vitro studies have demonstrated that the activity of BER enzymes are attenuated, though not completely inhibited, in nucleosome substrates. Several intrinsic properties of nucleosomes together with active cellular mechanisms, as introduced above, are capable of modulating DNA accessibility, revealing the complexity of the chromatin environment for BER.

This thesis investigates the initiation of BER by DNA glycosylases in the context of packaged DNA, and to which extent the packaging effects are influenced
by histone PTMs, the presence of histone tails, and histone variant incorporation. With these studies, we aim to acquire a global repair profile of different glycosylases and to understand the packaging effects of genomic DNA on the initiation of BER within the context of chromatin, with the long-term goal of providing insight into the balance between DNA damage, repair, and mutation signature of related diseases. Furthermore, we aim to advance the understanding of the molecular mechanisms of regulating BER by other cellular factors.
Figure 1.2 Examples of nucleobase lesions caused by oxidation, alkylation, deamination and hydrolysis. Moieties in red indicate the modification(s) to the primary nucleobases. Shown are 8oxoG, Tg (thymine glycol), DHU (5,6-dihydrouracil), 5hmU (hydroxymethyluracil), 5hC (5-hydroxycytosine), FapyG (4,6-diamino-5-formamidopyrimidine G), FapyA (4,6-diamino-5-formamidopyrimidine A), Gh (guanidinoxydantoin), Sp (spiroiminodiydantoin), 7mG (7-methylguanine), 3mG (3-methylguanine), 7mA (7-methyladenine), 3mA (3-methyladenine), εA, Hx (hypoxanthine), X (xanthine), and U. The AP site (abasic site) results from the hydrolysis of a glycosidic bond.
Figure 1.3 Short-patch sub-pathway of BER. (A) A glycosylase initiates the repair event by recognizing and excising the damaged nucleobase, creating an AP site. APE1 cleaves the DNA backbone 5’ to the AP site, generating 3’-OH and 5’-dRP termini. Pol β then converts the 5’-dRP to 5’-phosphate (5’-P) and inserts the correct nucleotide at the 3’-OH. In the last step, a ligase seals the nick in the DNA backbone to complete the repair process. (B) Cleavage of the glycosidic bond of an 8oxoG lesion by OGG1.
Figure 1.4 DNA packaging and NCP structure. (A) NCPs, connected by linker DNA, are compacted in the presence of histone H1 to form chromatin, which is further condensed into a chromosome. (B) A representative crystal structure of an NCP (PDB 1kx5). 147 bp of DNA is wrapped around a histone octamer core, which consists of two copies each of the four histones: H2A (yellow), H2B (red), H3 (blue), and H4 (green). The dyad axis is shown as a dashed line. (C) A side view of the NCP. One representative unstructured histone tail that protrudes from the NCP is indicated. (D) The rotational positions of three nucleobases, with the outward-facing, midway-facing and inward-facing positions highlighted in blue, purple and red, respectively. Histones are colored gray and only one DNA strand is shown for clarity.
Figure 1.5 Transient unwrapping of NCP. The wrapping and unwrapping of nucleosomal DNA create (A) the wrapped state and (B) the unwrapped state of the NCP.

Figure 1.6 NCP octasome and hexasome species. Histones are shown in the surface mode with H2A, H2B, H3 and H4 colored yellow, red, blue and green, respectively. (A) In the octasome, 147 bp of DNA is wrapped around a histone octamer core (PDB 1kx5). (B) The hexasome contains one copy of the H2A-H2B dimer and a (H3-H4)$_2$ tetramer, and only $\sim$110 bp of DNA are tightly associated with the histone hexamer core. The hexasome model was constructed by removing one H2A-H2B dimer from the octasome structure (PDB 1kx5) using PyMOL.
1.8 References


21. Bilotti K, Kennedy EE, Li C, Delaney S (2017) Human OGG1 activity in nucleosomes is facilitated by transient unwrapping of DNA and is influenced by
the local histone environment. *DNA Repair* 59:1–8.


119. Duarte LF, et al. (2014) Histone H3.3 and its proteolytically processed form


Chapter 2

Human OGG1 activity in nucleosomes is facilitated by transient unwrapping of DNA and is influenced by the local histone environment

2.1 Abstract

If unrepaired, damage to genomic DNA can cause mutations and/or be cytotoxic. Single base lesions are repaired via the base excision repair (BER) pathway. The first step in BER is the recognition and removal of the nucleobase lesion by a glycosylase enzyme. For example, human oxoguanine glycosylase 1 (OGG1) is responsible for removal of the prototypic oxidatively damaged nucleobase, 8-oxo-7,8-dihydroguanine (8oxoG). To date, most studies of glycosylases have used free duplex DNA substrates. However, cellular DNA is packaged as repeating nucleosome units, with 145 base pair segments of DNA wrapped around histone protein octamers. Previous studies revealed inhibition of OGG1 at the nucleosome dyad axis and in the absence of chromatin remodelers. In this study, we reveal that even in the absence of chromatin remodelers or external cofactors, OGG1 can initiate BER at positions off the dyad axis and that this activity is facilitated by spontaneous and transient unwrapping of DNA from the histones. Additionally, we find that solution accessibility as determined by hydroxyl radical footprinting is not fully predictive of glycosylase activity and that histone tails can suppress OGG1 activity. We therefore suggest that local nuances in the nucleosome environment and histone-DNA interactions can impact glycosylase activity.

2.2 Introduction

Genomic DNA is under constant threat of damage from free radicals, radiation, and environmental toxins (1). Persistent chemical modification of a nucleobase can be cytotoxic and/or mutagenic, and is the underlying cause of aging and cancer (1).
Therefore, cells have developed a variety of essential repair pathways to remove lesions from DNA. Single base lesions are rectified via the base excision repair (BER) pathway, which is initiated by a glycosylase enzyme specific to the lesion. For example, the oxidation of guanine to 8-oxo-7,8-dihy-droguanine (8oxoG) is repaired in humans by oxoguanine glycosylase 1 (OGG1). Glycosylases bind the DNA and flip the lesion out of the base stack into the enzyme’s active site. Glycosylases share a common S_N1 type mechanism for cleavage of the glycosidic bond at the lesion, ultimately resulting in an abasic site (2–4).

The majority of studies to examine glycosylase activity have used free duplex DNA as the substrate. While these studies provide essential mechanistic information about glycosylases, cellular DNA packaged as chromatin presents a more complex substrate context for repair in vivo (5). The primary structural repeating unit of chromatin is the nucleosome. Individual nucleosomes, or nucleosome core particles (NCP), serve as a model system for the lowest-order element of DNA packaging. An NCP consists of 145-147 base pairs of duplex DNA wrapped around a protein core. The octameric protein core contains two copies each of the four histone proteins: H2A, H2B, H3, and H4. The histone proteins have unstructured lysine-rich tails which are subject to post-translational modification, such as acetylation (6). The NCP contains a dyad axis of symmetry aligned with the central base pair of the wrapped DNA. The translational position of a base refers to its distance from the central base pair. Further, the rotational position of a base describes its orientation relative to the histone core. Outward facing bases are solution accessible while bases facing inward to the histone core are sequestered.

Several factors have been proposed to impact the initiation of repair by glycosylases in packaged DNA substrates, including the steric bulk of the histone proteins and the specific position of the lesion. A recent review by Odell et al. summarized
the studies to date and proposed a set of general tenets for glycosylase activity in NCPs (5). First, spontaneous unwrapping of nucleosomal DNA transiently exposes sites that are normally occluded. Ensemble FRET measurements (7–9) and single molecule FRET techniques (10, 11) have quantified the rates and extent of DNA end unwrapping. This dynamic motion ultimately results in increased accessibility of bases closer to the ends of DNA compared to bases at the dyad axis. Early work using restriction enzymes showed increased accessibility of restriction sites closer to ends of nucleosomal DNA compared to restriction sites on the dyad axis (12, 13). This trend was also observed with the human glycosylase endonuclease III (NTH1), which had increased excision of thymine glycol at the ends of the DNA compared to on the dyad axis (14); interestingly, it was recently reported that human cells contain a factor that facilitates NTH1 activity in NCP substrates (15). Further, previous studies have revealed significant inhibition of several other glycosylases, including OGG1, acting on lesions positioned at the dyad axis (14, 16, 17). Notably, the abilities of uracil DNA glycosylase (UDG) and alkyladenine glycosylase to excise an outward facing lesion at the dyad axis were observed as exceptions (16). Finally, it has been proposed that lesions with outward rotational position are more easily repaired than inward facing lesions, though exceptions have been observed (14, 16, 18–24). Most importantly, the general applicability of these principles to all glycosylases and lesions has yet to be fully explored.

In the case of OGG1, inhibition has been observed in NCP substrates when the lesion is located at the dyad axis (16, 17, 25). Activity of OGG1 in NCP substrates has been observed upon addition of chromatin remodeling complexes (17, 25) and when the lesion is located in the linker region between two NCPs of a dinucleosome substrate (25). Given the known importance of transient DNA unwrapping, OGG1 activity with lesions of varied translational position warrants further exploration.
To investigate the role of DNA unwrapping on OGG1 activity, we created NCP substrates with single base lesions incorporated off the dyad axis. The lesions were positioned approximately 20 base pairs from the end of the DNA within the region that is known to show dynamic unwrapping. To further investigate how the rotational position of a lesion impacts glycosylase activity, we incorporated lesions in positions of varying solution accessibility. Finally, we investigated the impact the local histone environment on OGG1 activity via acetylation and omission of the histone H2B tail.

2.3 Materials and Methods

2.3.1 Oligonucleotide synthesis and purification.

Oligomers were synthesized on a MerMade 4 (BioAutomation) DNA synthesizer using phosphoramidite chemistry. All synthesis reagents and phosphoramidites were purchased from Glen Research. The final trityl group was retained for an initial round of HPLC purification (Dynamax Microsorb C18 column, 10 × 250 mm; A = acetonitrile, B = 30 mM ammonium acetate; 5:95 to 35:65 A:B over 30 min at 3.5 mL/min). Following this first round of purification, oligomers were incubated with 20% v/v aqueous acetic acid for 60 min at room temperature to remove the trityl group. For oligomers containing only canonical bases or uracil (U), HPLC at 90 °C was used for a second round of purification (Agilent PLRP-S column, 250 mm × 4.6 mm; A = 100 mM triethylammonium acetate [TEAA] in 5% aqueous MeCN, B = 100 mM TEAA in MeCN; 0:100 to 35:65 A:B over 35 min at 1mL/ min). For oligomers containing 8oxoG, incorporation of 8oxoG was carried out according to Glen Research protocols. The trityl group was removed on the synthesizer before purification.
and 2-mercaptoethanol was included during cleavage of the oligomer from the bead. Anion-exchange HPLC was used to purify the 8oxoG-containing oligomers (Dionex DNAPac PA100 anion-exchange column; A = 10% acetonitrile, B = 0.8 M ammonium acetate in 10% acetonitrile; 70:30 to 0:100 A:B over 35 min at 1 mL/min). Collected peaks were desalted by buffer exchange using centrifugal concentrators (Sartorius Vivaspin Turbo 15, 5 kDa MWCO).

2.3.2 Ligation of 145 mer oligonucleotides.

We synthesized the DNA strands in shorter pieces for assembly into full length 145 mer via enzymatic ligation (see §2.7 for DNA sequences and ligation details). The short oligomers were phosphorylated with T4 polynucleotide kinase (New England Biolabs) with an additional 2 mM ATP in the reaction buffer. Phosphorylated oligomers were annealed to scaffold oligomers in a 1:1 ratio by heating to 90 °C for 5 min followed by cooling to 25 °C at a rate of 1 °C/min. Next, the oligomers were ligated for 3 hrs at ambient temperature using T4 Ligase (New England Biolabs). Full-length ligation products were purified by 8% denaturing PAGE (0.8 mm thickness). Electrospray ionization mass spectrometry was used to confirm the identity of the purified 145 mers. Oligonucleotide concentrations were determined by their absorbance at 260 nm using molar extinction coefficients calculated with OligoAnalyzer 3.1 (www.idtdna.com).

2.3.3 Glycosylase expression and purification.

His6-tagged OGG1 was recombinantly expressed in E. coli and purified as previously described (16, 26). SDS-PAGE analysis showed OGG1 purity to be > 98%. UDG was purchased from New England Biolabs. The total concentration of each enzyme was determined by Bradford assay with γ-globulin standards (Bio-Rad Laboratories).
2.3.4 Histone proteins preparation and octamer assembly.

Expression and purification of canonical *X. laevis* histone proteins and assembly of the histone octamer were performed according to the published method of Luger *et al.* (27, 28). Chemically acetylated H2B protein was prepared based on previously published methods (29, 30), which provides solution-accessible lysine residues as potential substrates. Briefly, 0.3 µmol of purified H2B was dissolved in 1 mL of freshly prepared 7 M urea solution with 500 mM ammonium acetate and 50 mM ammonium bicarbonate (pH 8.5). The mixture was kept on ice, and 1.5 µL of fresh 10.6 M acetic anhydride was added every 15 min. Concurrently, the pH of the reaction was constantly adjusted to 8.0-8.5 with ammonium hydroxide. After 60 min, the reaction was quenched with 100 µL of 1 M Tris-HCl before transferring to a dialysis device for overnight dialysis into water with 2 mM 2-mercaptoethanol. Electrospray ionization mass spectrometry was used to evaluate the addition of acetyl groups to H2B. The acetylated H2B protein was lyophilized and stored at 20 °C until used for octamer assembly. The tailless *X. laevis* H2B protein (residues 24-122) was purchased from The Histone Source (Colorado State University).

2.3.5 Nucleosome core particle (NCP) reconstitution.

NCPs were prepared by stepwise dialysis as reported previously (16, 31). Briefly, a Slide-a-Lyzer MINI dialysis device was equilibrated in buffer (10 mM Tris-HCl [pH 7.5], 1 mM EDTA, 1 mM dithiothreitol [DTT], 2 M NaCl) at 4 °C. Radiolabeled duplex (50 µL of 1 µM) was added to the dialysis device and allowed to incubate for 30 min before addition of histone octamer in a 1:1.05 molar ratio. The concentration of NaCl in the dialysis buffer was progressively lowered at 60 min intervals (1.2 M, 1.0 M, 0.6 M, 0 M). The final dialysis step was carried out for 3 hrs. Samples were filtered
to remove precipitate and NCP formation was confirmed by 7% native polyacrylamide (60:1 acrylamide:bisacrylamide; 0.25X TBE) gel electrophoresis (3 hrs at 150 V, 4 °C). Only NCP preparations with less than 5% free duplex DNA (as determined by native gel) were used in the experiments.

2.3.6 Hydroxyl radical footprinting.

Hydroxyl radical footprinting procedures were based on previously published methods (32, 33). Briefly, 5 pmol of NCPs containing $^{32}$P radiolabeled DNA were suspended in 52.5 µL buffer (10 mM Tris-HCl [pH 7.5], 1 mM EDTA) and mixed with 7.5 µL each 10 mM Fe(II)-EDTA, 10 mM sodium ascorbate, and 0.12% w/v aqueous hydrogen peroxide. The reaction was incubated at ambient temperature for 10 min before quenching with the addition of 50 µL 1 mM EDTA in 25% w/v glycerol. The sample was immediately loaded on a pre-running 7% native polyacrylamide (60:1 acrylamide:bisacrylamide, 0.25X TBE) and electrophoresed for 3 hrs at 150 V at 4 °C. The band containing NCPs was excised and NCPs were eluted into buffer (300 mM sodium acetate [pH 8.0], 1 mM EDTA) overnight. The resulting eluent was concentrated (Sartorius Vivaspin Turbo 15, 5kDa MWCO) and extracted twice against 25:24:1 phenol:chloroform:isoamyl alcohol. An ethanol precipitation was performed with the addition of 20 µL co-precipitation reagent (0.5 mg/ml tRNA in 300 mM sodium acetate [pH 8.0], 1 mM EDTA). Samples were dissolved in formamide and run on an 8% denaturing PAGE. The gel was dried and exposed before phosphorimaging. The bands were quantitated using SAFA gel analysis software (34).
2.3.7 Glycosylase kinetics experiments.

Kinetics experiments were based on previously published protocols (16). Briefly, radiolabeled, lesion-containing substrate (either duplex or NCP-incorporated) and glycosylase were prepared at 2X experimental concentration (final concentration 20 nM and 0.64 µM, respectively) in reaction buffer (20 mM Tris-HCl [pH 7.6], 25 mM NaCl, 75 mM KCl, 1 mM EDTA, 1 mM DTT, 200 µg/ml BSA). Following a temperature pre-equilibration at 37 °C for 2 min, equal volumes of substrate (8 µL) and glycosylase (8 µL) preparations were mixed for varying amounts of time before addition of 1 M NaOH quench (16 µL, final concentration 500 mM). For each time course a negative control sample (QC) was prepared by adding 1 M NaOH quench (16 µL) to substrate (8 µL) followed by addition of glycosylase (8 µL) before incubation at 37 °C for the duration of the longest time point. This control serves to reveal any pre-existing damage or incidental damage due to heating or sample work-up (generally less than 10%). Samples were heated to 90 °C for 2 min after addition of quench to induce a strand break at abasic sites. DNA was isolated from proteins by extraction with 25:24:1 phenol:chloroform:isoamyl alcohol before desalting by ethanol precipitation. Samples were electrophoresed on an 8% denaturing PAGE, imaged by phosphorimagery (BioRad Pharos FX) and quantitated by densitometry. The fraction of product at each time point, \( F_P(t) \), was determined by the formula

\[
F_P(t) = \frac{\delta_P(t)}{\delta_S(t) + \delta_P(t)},
\]

where \( \delta_S(t) \) and \( \delta_P(t) \) are the densities of the substrate and product bands, respectively, at time \( t \). The product yield, \( P(t) \), was corrected for pre-existing and incidental
substrate damage using the following formula:

\[ P(t) = \frac{F_P(t) - F_P(0)}{1 - F_P(0)} , \tag{2.2} \]

where \( F_P(0) \) is the fraction of product observed in the QC sample. The mean product yield from four independent trials (two separate NCP preparations with two trials each) was determined at each time point. The averaged data were fit to the modified first-order integrated rate law:

\[ [P(t)] = P(\infty)(1 - e^{-k_{obs}t}) , \tag{2.3} \]

or

\[ [P(t)] = [P1(\infty)](1 - e^{-k_{1,obs}t}) + [P2(\infty)](1 - e^{-k_{2,obs}t}) , \tag{2.4} \]

where \([P(t)]\) is the concentration of product at time \( t \), and \([P(\infty)]\) is the maximum concentration of product, using a nonlinear least squares regression (Kaleidagraph). Reaction rates, \( k_{obs} \), were reported from the fits. A two-tailed student’s \( t \) test was performed to obtain the \( p \) values for the tailless H2B NCPs and acetylated H2B NCPs in comparison to the canonical NCPs. We considered \( p < 0.05 \) to be significant.

### 2.3.8 Molecular modeling.

Molecular models were created in PyMOL (The PyMOL Molecular Graphics System, Version 1.74 Schrödinger, LLC.) to visualize OGG1 binding to the 601 NCP. Structures of OGG1-bound-NCPs were created by aligning the crystal structure of OGG1 bound to DNA (PDB ID: 1ebm) with the 601 NCP crystal structure (PDB ID: 3lz0) as described previously (16). The glycosylase surface was colored in PyMOL according to proximity to the histone core (regions within 5 Å of the octamer are yellow;
regions within 5-10 Å of the octamer are red; regions greater than 10 Å from the octamer are blue).

### 2.4 Results

#### 2.4.1 Formation of homogenous NCPs.

In this study, we evaluated the impact of lesion position on BER in packaged DNA. Our DNA substrates were based on the Widom 601 positioning sequence, which binds in a single translational and rotational position around the histone octamer (35). We used the Widom 601 NCP crystal structure (36) to guide our placement of DNA lesions off the dyad axis and in one of three rotational positions: out toward solution (OUT), approximately 90° from solution (MID), or in toward the histone core (IN) (Figure 2.1A and C). These OUT, MID, and IN positions correspond to locations 24, 21, and 19, respectively, which indicate the distance from the 5' end of the DNA strand. A single 8oxoG or U was incorporated at these sites during DNA synthesis (Figure 2.1D). The complementary 601 strand was modified such that 8oxoG was paired with C and U was paired with G to mimic a deaminated C:G base pair. *X. laevis* histone proteins were individually expressed in *E. coli*, purified, assembled into an octamer, and used to form NCPs following the salt dialysis method (31). Histone octamers containing either a tailless or acetylated H2B protein were also used to assemble NCPs. Figure 2.2 shows a representative native gel that indicates the formation of each NCP type.
Figure 2.1 Representation of the NCP structure and lesions used in this study. (A) Merged crystal structure of an NCP containing Widom 601 DNA and histone octamer with tails (PDB codes 3lz0 and 1kx5, respectively) with off-dyad lesion locations highlighted using PyMol. Bases in the lesion-containing strand are numbered starting from the 5′ end. The OUT, MID, and IN facing lesion is positioned at base 24, 21, and 19, respectively. (B) Side view of NCP. The H2B tail is highlighted in green and represented in surface mode. (C) Rotational positions of lesions (one DNA strand shown for simplicity). The location of the OUT, MID and IN facing lesion is highlighted in red, purple, and blue, respectively. (D) Lesions used in this study.
Figure 2.2 Representative native PAGE showing the formation of NCP. Native PAGE was used for evaluating the reconstitution of canonical, tailless H2B (tlH2B), and acetylated H2B (AcH2B) NCPs. Radiolabeled samples were loaded on a 1 mm thick 7% native gel (60:1 acrylamide:bisacrylamide, 0.25X TBE). The gel was run at 4 °C for 3 hrs at 150 V. Variable migration distances are observed for single strand, duplex, and NCP-incorporated DNA samples.

2.4.2 Hydroxyl radical footprinting defines rotational position of lesions.

Hydroxyl radical footprinting was used to establish the solution accessibility of each lesion position. The hydroxyl radical footprinting technique uses Fenton chemistry to generate hydroxyl radicals, which abstract a hydrogen atom from the DNA backbone to create a strand break (33). A sample of free duplex Widom 601 DNA treated with hydroxyl radicals shows an unbiased pattern of reactivity throughout the length of the DNA strand (Figure 2.3A, duplex lane). In comparison, DNA incorporated into an NCP shows an oscillating pattern of damage (Figure 2.3A, NCP lanes). There are regions of protection and damage corresponding to the DNA wrapping in toward the histone core and out toward solution, respectively. Notably, the oscillatory pattern persists both upstream and downstream of the lesion positions, indicating that while this region is known to be more dynamic (11), it is still in contact with the histone
core and does not react like free duplex DNA with hydroxyl radical treatment. The integrated band area at each base position, and the relative reactivity of lesion positions is shown in Figure 2.3B (34). Indeed, the relative hydroxyl radical reactivity at the selected lesion positions correlates with predicted solution accessibility based on the Widom 601 NCP crystal structure (ie. OUT > MID > IN).

Further, hydroxyl radical footprinting was used to evaluate the impact of changes to histone proteins on the solution accessibility of the DNA bases. NCPs prepared using a histone octamer containing either tailless H2B or acetylated H2B histone proteins showed no systematic differences in the oscillating pattern or the relative solution accessibility of the studied OUT, MID, and IN positions when compared to the canonical histone octamer (Figure 2.3).

2.4.3 Glycosylase activity in NCPs with off-dyad lesions.

Kinetics experiments were performed to test the activity of glycosylase enzymes on NCP substrates. NCP substrates containing a single 8oxoG or U lesion were incubated at 37 °C with OGG1 or UDG before quenching with NaOH. Formation of product over time was monitored on denaturing polyacrylamide gels (Figure 2.4). In these experiments, a 32-fold excess of enzyme over substrate was used to ensure single-turnover conditions. We have previously shown that further increases in the excess of enzyme do not yield an increased rate of glycosylase activity in NCP substrates (16). The observed rate ($k_{obs}$) therefore reflects the slowest step up to or including the chemistry step of glycosidic bond cleavage.

Control experiments were performed in which free duplex DNA substrates containing a single 8oxoG or U lesion (the OUT facing lesion position is shown as a representative example) were converted to 87% and 99% product, respectively (Figure
Figure 2.3 Hydroxyl radical footprinting (HRF) of DNA in an NCP. (A) The HRF experiment was analyzed by denaturing PAGE. The flanking “A+G” lanes display a ladder of the 601 DNA sequence created using Maxam Gilbert reactions. The “NT” lane shows a no treatment control. The “duplex” lane shows unbiased damage of hydroxyl radical treatment on free duplex DNA. Hydroxyl radical treatment of canonical (“can. NCP”), tailless H2B (“tlH2B NCP”), and acetylated H2B (“AcH2B NCP”) NCPs reveals an oscillating pattern of damage. The location of OUT (red), MID (purple), and IN (blue) lesions in this study are highlighted. (B) Quantitation of band density of the footprinting using SAFA software to show relative solvent accessibility. The vertical lines highlight the location of each lesion position.
Figure 2.4 A representative denaturing PAGE showing the excision of U by UDG under single-turnover condition. The conversion of the lesion containing NCP substrates (S) to nicked product (P) is observed with increasing time. UDG acting on duplex (“Dup” lane) is included to verify the activity of UDG. The “-E” lane is a negative control prepared by incubating the NCP substrate in the absence of UDG at 37 °C for the duration of the longest time point. The “QC” lane reveals any pre-existing damage or incidental damage, and was prepared by adding 1 M NaOH quench (16 µL) to substrate (8 µL) followed by the immediate addition of UDG (8 µL) before incubation at 37 °C for the duration of the longest time point.

2.5. The data were fit to a single exponential, and the differences in $k_{\text{obs}}$ and product formation for free duplex DNA substrates of each lesion/glycosylase pair varied minimally with lesion position (data not shown).

For NCP substrates, we previously showed complete inhibition of OGG1 activity on lesions located at the dyad axis, regardless of rotational position (16). Interestingly, in the current work, we observe that activity of OGG1 is restored on NCP substrates when the lesions are located off the dyad axis (Figure 2.5A). Surprisingly, however, the product yield of these kinetic time courses does not correlate with the solvent accessibility of the lesions as determined by hydroxyl radical footprinting. The most product formation (93%) was observed at the MID facing lesion site and 49% product formation was observed for the IN facing lesion site. The most unanticipated result was the relative inhibition of OGG1 at the most solvent accessible lesion tested in this experiment (OUT), which reached a maximum of 38% product (Figure 2.5A, Table 2.1). In contrast to the free duplex DNA substrates, the data for all
Figure 2.5 Single-turnover kinetic results for OGG1 and UDG in NCP substrates. Kinetic time courses were performed to evaluate the activity of (A) OGG1 or (B) UDG on either free duplex DNA (black circle) or NCP substrates with off-dyad lesions. The NCP substrates contained lesions of varying rotational position: OUT (red diamond), MID (purple triangle) or IN (blue square). Experiments were conducted using 20 nM NCP substrate and 0.64 μM glycosylase in 20 mM Tris-HCl (pH 7.6), 25 mM NaCl, 75 mM KCl, 1 mM EDTA, 1 mM DTT, 200 μg/ml BSA. Data for free duplex DNA substrates were fit to a single exponential equation; all NCP data were fit to a double exponential equation. Error bars represent the standard error (n=4).
Table 2.1. $k_{\text{obs}}$ values determined for duplex DNA and NCP substrates.

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>NCP</th>
<th>Substrate</th>
<th>$k_{\text{obs}}$/min (%product)</th>
</tr>
</thead>
<tbody>
<tr>
<td>OGG1</td>
<td>N/A</td>
<td>8oxoG duplex</td>
<td>2.6 ± 0.2 (87%)</td>
</tr>
<tr>
<td></td>
<td>canonical</td>
<td>8oxoG\textsuperscript{OUT}</td>
<td>1.4 ± 0.4 (23%), 0.03 ± 0.03 (15%)</td>
</tr>
<tr>
<td></td>
<td>canonical</td>
<td>8oxoG\textsuperscript{MID}</td>
<td>2.0 ± 0.4 (63%), 0.06 ± 0.04 (30%)</td>
</tr>
<tr>
<td></td>
<td>canonical</td>
<td>8oxoG\textsuperscript{IN}</td>
<td>0.8 ± 0.1 (25%), N.D.\textsuperscript{a} (≥ 24%)</td>
</tr>
<tr>
<td></td>
<td>tailless H2B</td>
<td>8oxoG\textsuperscript{OUT}</td>
<td>2.3 ± 0.4 (28%), 0.05 ± 0.02 (18%)</td>
</tr>
<tr>
<td></td>
<td>acetylated H2B</td>
<td>8oxoG\textsuperscript{OUT}</td>
<td>4.7 ± 1.4 (23%), 0.11 ± 0.03 (33%)</td>
</tr>
<tr>
<td>UDG</td>
<td>N/A</td>
<td>U duplex</td>
<td>12.6 ± 1.6 (99%)</td>
</tr>
<tr>
<td></td>
<td>canonical</td>
<td>U\textsuperscript{OUT}</td>
<td>37.7 ± 33.4 (29%), 0.3 ± 0.05 (67%)</td>
</tr>
<tr>
<td></td>
<td>canonical</td>
<td>U\textsuperscript{MID}</td>
<td>20.1 ± 5.3 (52%), 3.3 ± 0.6 (46%)</td>
</tr>
<tr>
<td></td>
<td>canonical</td>
<td>U\textsuperscript{IN}</td>
<td>42.3 ± 19.9 (34%), 0.3 ± 0.04 (33%)</td>
</tr>
</tbody>
</table>

\textsuperscript{a}N.D. indicates an undetermined $k_{\text{obs}}$ value for this kinetic phase.

NCP substrates were best fit to a double exponential model, reflecting a faster kinetic phase and a slower kinetic phase (Table 2.1). While the product yield varies between different rotationally positioned substrates, the $k_{\text{obs}}$ values are the same within error.

The unexpected pattern of glycosylase activity and solvent accessibility was also manifested for another glycosylase, UDG, acting on the off-dyad lesions. However, for UDG, rather than observing differences in the amount of product, we observed differences in $k_{\text{obs}}$; while full product conversion was achieved for both the OUT and MID positions, U at the MID position was excised approximately 10-fold faster than at the OUT position (Figure 2.5B, Table 2.1).
2.4.4 Glycosylase activity in NCPs with H2B-modified octamers.

Upon observing the disaccord between the solvent accessibility of the lesions and the resulting glycosylase activity, we modeled the Widom 601 DNA wrapped around a histone core containing histone tails (PDB ID 1kx5), which are present in our *in vitro* experiments. In this model we observed that our off-dyad lesions are near the H2B tail that protrudes between the superhelices (Figure 2.1B). We hypothesized that interactions between the histone tail and the incoming glycosylase could explain the suppressed product yield of OGG1 at the OUT lesion position as well as the slower $k_{\text{obs}}$ for UDG.

To test the impact of the nearby H2B tail on OGG1 activity, we created two modified histone octamers. First, we created NCPs using an octamer with chemically acetylated H2B proteins, where addition of an acetyl group neutralizes the positive charge of lysine. Electrospray ionization mass spectrometry was used to confirm and evaluate the acetylation of H2B (Figure 2.6). In order to further test the steric impact of the H2B tail, we created NCPs using an octamer with tailless H2B proteins. We tested the activity of OGG1 on these H2B-modified NCPs with an OUT facing 8oxoG and compared product formation to the canonical NCP substrate. While the tailless H2B NCP does not show a statistically significant increase in product formation, there is an increase in product with the acetylated H2B substrate ($p = 0.0016$) (Figure 2.7).

As a comparison point, we tested the impact of the H2B tail on the activity of UDG on off-dyad U lesions incorporated into the tailless H2B NCP. When the U was placed at the OUT position in the tailless H2B NCP, the observed rate was comparable to that seen in the canonical NCP (Figure 2.8A). Similarly, the observed rate and the amount of product accumulation of U removal at the IN position were
Figure 2.6 Electrospray ionization mass spectra of acetylated H2B. (A) Mass spectrometry confirms H2B acetylation. The peak at 13914.29 amu corresponds to the addition of 10 acetyl groups. (B) A zoomed in view of the mass spectrum in (A) reveals a distribution in the number of acetyl group added. Peaks corresponding to the addition of 4-14 acetyl groups were observed.

Figure 2.7 Single-turnover kinetic results for OGG1 on 8oxoG^{OUT} in NCPs containing modified H2B. The octamer used to create the NCPs was varied to include a canonical octamer (diamond, solid line), tailless H2B octamer (inverted triangle, short dashed line) or acetylated H2B octamer (triangle, long dashed line). Experiments were conducted using 20 nM NCP substrate and 0.64 µM total OGG1 in 20 mM Tris-HCl (pH 7.6), 25 mM NaCl, 75 mM KCl, 1 mM EDTA, 1 mM DTT, 200 µg/ml BSA. All data were fit to a double exponential equation. Error bars represent the standard error (n=4).
Figure 2.8 Single-turnover kinetics for the removal of off-dyad U lesions by UDG in tailless H2B NCPs. The U was placed in an NCP facing (A) OUT or (B) IN, with the data for tailless H2B NCP and canonical NCP shown in dashed lines and solid lines, respectively. All NCP data were fit to a double exponential equation. Error bars for tailless H2B NCP data are the standard error (n=2). Canonical NCP data for each position is duplicated from Figure 2.5B for comparison purposes.

also comparable to those seen in the canonical NCP (Figure 2.8B).

2.5 Discussion

To investigate the role of DNA unwrapping on glycosylase activity, we created NCP substrates with single base lesions located approximately 20 base pairs from the end of the DNA. In general, activity of the tested enzymes increases when the lesions are moved closer to the DNA ends, relative to when the lesions are at the dyad. While OGG1 was completely inhibited at the dyad axis regardless of the rotational position of the lesion (16), product formation is observed at all of the off dyad positions studied here.

While the data for free duplex DNA substrates fit well to a single exponential
model, the kinetic data for the NCP substrates are more complex. For the off-dyad lesions, double exponential models were required to fit the data for all lesion positions and glycosylases. We attribute the biphasic kinetics to the existence of two distinct substrate populations. Each NCP substrate shows a fast phase, which we attribute to a configuration that can be processed directly by the glycosylase. The slow kinetic phase for NCP substrates is generally two orders of magnitude slower than the fast phase. We attribute the slower phase to a population of NCPs that require a rate-limiting conformational change to adopt a structure permissive for repair. This conformational change is consistent with the transient unwrapping of DNA to expose the off-dyad lesion sites, and is not observed in experiments with lesions positioned on the dyad (16). Previously, the Stivers group reported multiphasic kinetics in experiments with UDG excising U at sites off the dyad (31).

As no crystal structures of glycosylase-bound NCPs have been reported, we used molecular modeling to analyze the interaction of the glycosylase and the NCP (Figure 2.9). To generate these models, we merged the crystal structure of the 601 NCP with the crystal structure of DNA-bound OGG1. We chose to model the glycosylase-bound NCP using the 601 NCP crystal structure lacking histone tails (36) in order to avoid assigning static positions to the unstructured histone tails. The surface of the glycosylase was colored according to its proximity to the histone core as follows: regions within 5 Å of the octamer are yellow; regions within 5-10 Å of the octamer are red; regions greater than 10 Å away from the octamer are blue.

Upon modeling OGG1 bound to the NCP at the outward facing off-dyad lesion, we found the region near the enzyme’s active site to have several amino acid residues within 5 Å of the octamer and an even larger region between 5 and 10 Å of the octamer. This region represents a significant steric clash between the histone core and OGG1, which we might predict would prevent enzyme binding and, ultimately,
excision of the lesion. However, 38% product formation is observed for OGG1 at this off-dyad lesion position. The formation of product despite the potential for steric clash in a static model at this position emphasizes the contribution of DNA unwrapping and dynamics to OGG1 accessibility in an NCP.

There are conflicting reports on the correlation between solution accessibility of a lesion and enzyme activity in an NCP substrate (14, 16, 18–24). We found that for OGG1, the product yield for NCP substrates did not correlate with the solution accessibility of the lesions as determined by hydroxyl radical footprinting. Furthermore, while full product conversion was observed for both the OUT and MID position with UDG, the rate of excision at the MID position was ~10 times faster than at the OUT position. It is clear that solution accessibility as determined by hydroxyl radical footprinting is not able to predict all chemistry within an NCP. We therefore suggest that local nuances in the NCP environment and histone-DNA interaction have a significant impact on glycosylase activity and may have a more influential role than
solution accessibility at certain positions in the NCP. Other groups who have observed complex kinetics of UDG in NCP substrates have come to similar conclusions (31).

We created a tailless H2B version of the histone octamer to test the influence of completely removing the histone tail on glycosylase activity. We found that for OGG1, removing the H2B tails does not lead to a statistically significant increase in product formation. While we therefore conclude that the inhibition of OGG1 is not completely derived from steric interference of the histone tail, the phenomenon of histone tail clipping \textit{in vivo} by endopeptidases is an area of recent interest and has been observed in all histone proteins in several organisms (37). It has been shown that histones lacking N-terminal tails allow for increased accessibility of nucleosomal DNA to transcription factors via increased unwrapping of DNA from the NCP (38–40). Other BER enzymes have also shown varying levels of inhibition in NCP substrates due to the presence of histone tails (41, 42).

Histone proteins are also targets for post-translational modification, particularly acetylation at the positively charged lysine residues in the N-terminal tails. In this study, NCPs with acetylated H2B proteins showed an increase in product formation when compared to unmodified NCPs. While the H2B tail associates with the nearby DNA in unmodified NCPs, it has been shown that acetylation causes release of H2B tails from the DNA (43). Further, FRET studies have shown that acetylation of histones facilitates DNA unwrapping (44, 45). We therefore attribute the significant increase in product formation to the neutralization of the lysine residues and resulting relaxation of the NCP structure. However, the incomplete recovery of product formation in acetylated H2B NCPs suggests that electrostatic interference of the lysine residues on H2B is not the only contributing factor to OGG1 inhibition at the OUT lesion position.
2.6 Conclusion

In summary, we demonstrate that OGG1 can excise 8oxoG from nucleosomal DNA in the absence of external cofactors or chromatin remodelers and that this activity is modulated by transient DNA unwrapping. We further demonstrate the complexity of the NCP environment, as we observe a disparity between solution accessibility and enzyme activity. Increased activity of OGG1 in response to acetylation of H2B reveals intricacies in the NCP environment due to the electrostatic influence of the histone tails. A more complete understanding of BER in NCP substrates will require additional studies on the impact of histone modifications, chromatin remodeling complexes, and other cellular factors.

2.7 Supporting information

DNA sequences used in the present work (46)

The 145 mer oligos were synthesized as three component oligos. Each strand was divided into a 45 mer, a 40 mer, and a 60 mer, as indicated by the vertical lines. The component oligos were assembled for ligation using the scaffolds listed. X indicates the location of the lesion (8oxoG or U) in the lesion strand, and Y indicates the location of the base opposite the lesion in the complement strand (C opposite 8oxoG and G opposite U).

145 mer lesion strand

5’- ATC AGA ATC CCG GTG CCG X^{19}G^{21} CCX^{24} CTC AAT TGG TCG TAG ACA GCT | CTA GCA CCG CTT AAA CGC ACG TAC GCG CTG TCC CCC GCG T | TT TAA CCG CCA AGG GGA TTA CTC CCT AGT CTC CAG GCA
CGT GTC AGA TAT ATA CAT CGA T -3'

scaffold 1
5'- TTT AAG CGG TGC TAG AGC TGT CTA CGA CCA -3'

scaffold 2
5'- CCC TTG GCG GTT AAA ACG CGG GGG ACA GCG -3'

145 mer complement strand
5'- ATC GAT GTA TAT ATC TGA CAC GTG CCT GGA GAC TAG GGA GAT ATC | CCC TTG GCG GTT AAA ACG CGG GGG ACA GCG CGT ACG TGC G | TT TAA GCG GTG CTA GAG CTG TCT ACG ACC AAT TGA GYG GYC YCG GCA CCG GGA TTC TGA T -3'

scaffold 3
5'- TTT AAC CGC CAA GGG GAT TAC TCC CTA GTC -3'

scaffold 4
5'- CTA GCA CCG CTT AAA CGC ACG TAC GCG CTG -3'
2.8 References


Chapter 3

Histone H2A variants enhance the initiation of base excision repair in nucleosomes

3.1 Abstract

Substituting histone variants for their canonical counterparts can profoundly alter chromatin structure, thereby impacting multiple biological processes. Here, we investigate the influence of histone variants from the H2A family on excision of uracil (U) by the base excision repair (BER) enzymes uracil DNA glycosylase (UDG) and single-strand selective monofunctional uracil DNA glycosylase (SMUG1). Using a DNA population with globally distributed U:G bp, enhanced excision is observed in H2A.Z and macroH2A-containing nucleosome core particles (NCPs). The U with reduced solution accessibility exhibit limited UDG activity in canonical NCPs but are more readily excised in variant NCPs, reflecting the ability of these variants to facilitate excision at sites that are otherwise poorly repaired. We also find that U with the largest increase in excision in variant NCPs are clustered in regions with differential structural features between the variants and canonical H2A. Within 35-40 bp from the DNA terminus in macroH2A NCPs, the activities of both glycosylases are comparable to that on free duplex. We show that this high level of activity results from two distinct species within the macroH2A NCP ensemble: octasomes and hexasomes. These observations reveal potential functions for H2A variants in promoting BER and preventing mutagenesis within the context of chromatin.

3.2 Introduction

As introduced in §1.2, cellular DNA is constantly exposed to damaging agents from both endogenous and exogenous sources, and therefore DNA repair is vital for maintaining genomic integrity and preventing mutagenesis (1). Detrimental chemical modification of nucleobases is repaired via the base excision repair (BER) pathway. A
DNA glycosylase enzyme initiates BER by recognizing its lesion substrate and cleaving the glycosidic bond, generating an abasic site, and the repair event is completed by downstream BER enzymes (2, 3).

Uracil (U) is one of the most prevalent nucleobase lesions in DNA (4). It can result from either misincorporation of dUMP opposite A during replication or spontaneous cytosine deamination, with the latter giving rise to U:G base pairs (bp). Spontaneous cytosine deamination is estimated to occur at a rate of 100-500 events per cell per day (4). It was recently reported that U accumulation in CpG sites arises primarily from cytosine deamination in mice lacking the two major glycosylases responsible for excising U from U:G bp (5–7), uracil DNA glycosylase (UDG) and single-strand selective monofunctional uracil DNA glycosylase (SMUG1) (8). If left unrepaired, U:G bp lead to C to T substitutions, which are the most common mutational signatures in human cancer (9). Compared with UDG, SMUG1 excises U opposite G less efficiently and has been suggested as a backup for UDG but with broader substrate specificity (6, 10). It is known that the highest levels of UDG are present during S phase and that UDG has both mitochondrial and nuclear isoforms, while the level of SMUG1 is not dependent on the cell cycle and is found only in the nucleus (11). These differences suggest distinct spatiotemporal roles for these two glycosylases.

Tremendous effort has been devoted to defining glycosylase activity using duplex DNA as the substrate (i.e., free in solution and with no DNA binding factors present). However, genomic DNA in eukaryotes is packaged into chromatin, with nucleosome core particles (NCPs) as the fundamental repeating unit (Figure 1.4A). An NCP consists of 145-147 bp of duplex DNA wrapped ~1.7 times around an octameric histone protein core, which is comprised of two H2A-H2B dimers and an (H3-H4)2 tetramer. Within an NCP, there exists a two-fold axis of pseudosymmetry known as
the dyad axis (12). The position of a nucleobase in an NCP is typically described in terms of its translational position, the displacement from the dyad axis, and its rotational position, the orientation inward towards the protein or outward towards the solution (Figure 1.4B, C, and D).

Due to the interaction with histones, DNA in an NCP is structurally and dynamically constrained, posing challenges to DNA-processing enzymes associated with multiple cellular processes including transcription and repair. Nucleobases facing outward are highly solution-accessible while those facing inward toward the histones are occluded and virtually solution-inaccessible. Nevertheless, the transient and spontaneous unwrapping of nucleosomal DNA can expose occluded nucleobases near the DNA entry/exit region (13, 14). Several other mechanisms can actively modulate DNA accessibility in cells, including histone post-translational modifications (PTMs), chromatin remodeling, and the incorporation of histone variants into chromatin.

Histone variants are nonallelic isoforms of canonical histones that exhibit varied degrees of divergence in amino acid sequence relative to their conventional counterparts. Replacement of canonical histones with variants can confer distinct structures and dynamics and, in turn, specialized functions to variant-containing NCPs. The H2A family contains the largest number of variants. Three structural segments of these H2A variants are essential for diversifying intra- and inter-nucleosomal associations (15, 16): the docking domain residing in the C-terminus, the L1 loop, and the acidic patch. One of the most extensively studied H2A variants is H2A.Z, which is found in almost all eukaryotes. Although H2A.Z is only ~60% identical to canonical H2A, it is highly conserved among species (15), indicating its unique and significant function. Indeed, H2A.Z has been implicated in a multitude of biological events including transcription, double-strand break (DSB) repair, nucleotide excision repair (NER), heterochromatin silencing, chromosome segregation, and progression
through the cell cycle (17–19). MacroH2A is another H2A variant of particular interest. It possesses an H2A-like histone domain connected via a lysine-rich linker to a C-terminal macro domain. In addition to inactivating X chromosomes (20) and orchestrating gene expression (21, 22), macroH2A has been shown to function in DSB DNA damage response (23). Based on the fact that H2A.Z and macroH2A are involved in other DNA repair pathways, we questioned whether these two variants play a role in BER as well.

In this work, we evaluated the global profile of U excision by UDG and SMUG1 in the context of packaged DNA, using a population of NCPs that contain U lesions with a wide variety of translational and rotational positions. By substituting canonical H2A with H2A.Z or macroH2A, we investigated the impact of these two H2A variants on the initiation of BER. Further, we performed single-turnover kinetics experiments to determine UDG and SMUG1 efficiency on NCPs containing these histone variants.

### 3.3 Materials and Methods

#### 3.3.1 Oligonucleotide synthesis and purification

The full sequences of all DNA strands used in this work are shown in §3.6. The 145 mer oligonucleotide DNA containing global U (LS\text{U\text{global}}) was synthesized on a MerMade 4 (BioAutomation) DNA synthesizer using phosphoramidite chemistry. All phosphoramidites and reagents for synthesis were purchased from Glen Research. Similar to a recent report (24), a mixture of C/U phosphoramidites was used during synthesis with the molar ratio determined by a Poisson distribution (\(\lambda=0.355\)) to ensure that 95\% of the DNA contained at most one U per strand. The DNA was
purified by 8% denaturing polyacrylamide gel electrophoresis (PAGE). To confirm the global U incorporation, the single-stranded DNA was 5'-radiolabeled with $^{32}$P by T4 kinase (New England Biolabs) and incubated with UDG at 37 °C for 30 min, followed by NaOH-catalyzed strand cleavage at resulting abasic sites at 90 °C for 2 min. Starting from the 5'-end of the strand, nucleobases are numbered from 1 to 145. The 145 mer with a single U at site 110 (LS$_{U110}$) and the 145 mer complement strand (CS) (which serves for both LS$_{U\text{global}}$ and LS$_{U110}$) were prepared via ligating shorter component strands using T4 ligase (New England Biolabs) in the presence of scaffold DNA strands (Figure 3.15). Component and scaffold strands were synthesized with the final trityl group retained for reversed-phase HPLC purification (Dynamax Microsorb C18 column, 250 × 10 mm; A = acetonitrile [MeCN], B = 30 mM NH$_4$OAc; 5:95 to 35:65 A:B over 30 min at 3.5 mL/min). The trityl group was then removed by incubation in 20% v/v aqueous glacial acetic acid for 1 h at ambient temperature, followed by a second round of HPLC purification at 90 °C (Agilent PLRP-S column, 250 × 4.6 mm; A = 100 mM triethylammonium acetate [TEAA] in 5% aqueous MeCN, B = 100 mM TEAA in MeCN; 0:100 to 15:85 A:B over 35 min, 15:85 to 35:65 A:B over 5 min at 1 mL/min). Electrospray ionization mass spectrometry was applied to confirm the identity of component and scaffold strands. The ligated LS$_{U110}$ and CS were purified by 8% denaturing PAGE. The two single-stranded internal standards, which serve as a reference for normalizing band quantification in experiments using global U-containing substrates, were designed as a 30 mer (IS$_{30}$) and a 98 mer (IS$_{98}$) such that they would not co-migrate with any DNA fragments resulting from U excision. They were synthesized with the final trityl group removed and purified by 12% and 8% denaturing PAGE separately.
3.3.2 Histone preparation, NCP reconstitution and heat-shifting assay

Canonical histones from *X. laevis* and the histone domain of human macroH2A (aa 1-121) were recombinantly expressed and purified according to published protocols (25). Human H2A.Z was purchased from the Histone Source (Colorado State University). Canonical H2A, H2A.Z and macroH2A containing octamers were individually assembled (25), and the corresponding NCPs were then reconstituted via salt-gradient dialysis as previously described (26, 27). Briefly, LS$_{U_{\text{global}}}$ or LS$_{U_{110}}$ was 5’-radiolabeled and annealed to CS. 1 µM duplex DNA was then mixed with an equimolar amount of canonical, H2A.Z or macroH2A octamers to initiate NCP reconstitution at 4 °C in buffer containing 2 M NaCl, 10 mM Tris-HCl (pH 7.5), 1 mM EDTA, 1 mM dithiothreitol (DTT) and 500 µg/mL BSA. The concentration of NaCl was reduced stepwise at 1 h interval (1.2 M, 1.0 M, 0.6 M and 0 M) via dialysis, and the final dialysis into 0 M was carried out for 3 h. After reconstitution, samples were filtered by centrifugation using a Spin-X Centrifuge Tube Filter (0.22 µm, Corning Incorporated) to remove precipitates, and the purity of NCPs was evaluated by 7% native PAGE (60:1 acrylamide : bisacrylamide; 3 h at 150 V, 4 °C; 0.25X TBE). Only NCPs that exhibited high purity (with < 2% duplex DNA that is not incorporated into NCPs) were used in subsequent experiments. In the heat-shifting assay, NCPs were incubated at 37 °C or 55 °C for 1 h and subjected to 7% native PAGE immediately at 4 °C.

3.3.3 Global assessment of UDG and SMUG1 activity

Similar to other studies that examined U excision from NCPs (28, 29), we used *E. coli* UDG that is 73% similar to human UNG and has the active site completely conserved (30). Furthermore, comparison of the crystal structures of *E. coli* UDG
and human UNG shows root mean square deviation of < 1Å when Cα are aligned. The overall shapes of these two orthologs are therefore very similar (31). *E. coli* UDG and human SMUG1 were purchased from New England Biolabs, and the total concentration of each glycosylase was determined by Bradford assay using γ-globulin standards (Bio-Rad Laboratories). To assess the activity of UDG and SMUG1, 0.5 pmol substrate (either duplex DNA or NCPs containing global U lesions) and the two 5'-radiolabeled internal standards were mixed with 10 pmol UDG or SMUG1 in a total volume of 20 µL of the reaction buffer (20 mM Tris-HCl [pH 7.6], 50 mM NaCl, 150 mM KCl, 1 mM EDTA, 1 mM DTT, 200 µg/mL BSA). The glycosylase reaction was performed at 37 °C for 60 min. Meanwhile, a negative control sample (-E) was incubated at 37 °C for 60 min, which included 0.5 pmol substrate and the two internal standards of the same amounts as those in the glycosylase-treated sample in a total volume of 20 µL of the reaction buffer. The negative control would reveal any pre-existing damage or incidental damage of the substrate before and during the reaction. After the 60-min incubation, 20 µL of 1 M NaOH were added to all samples, followed by heating at 90 °C for 2 min. NaOH can immediately deactivate the glycosylase for efficient reaction quenching and catalyze strand break at an abasic site. DNA fragments were next isolated from proteins by extraction against 25:24:1 phenol:chloroform:isoamyl alcohol (PCI). Ethanol precipitation was carried out by supplementing the resulting aqueous phase with 40 µL co-precipitation reagent (0.3 M NaOAc, 1 mM EDTA, 0.5 mg/mL tRNA) and 600 µL ethanol and incubating at -20 °C for overnight. The sample was resuspended in 50% v/v formamide, split in half, and half loaded onto a 10% denaturing PAGE gel (1.5 h at 80 W, 1X TBE) for resolving bands from nucleobase 9 to the dyad axis and half loaded onto an 8% gel (3 h at 80 W, 1X TBE) for resolving bands from the dyad axis to nucleobase 129. Gels were imaged by phosphorimagery (Bio-Rad PharosFX), and bands were quantified
using SAFA software (32). To correct for any slight difference in the amount of each sample loaded on the gel, quantified band intensities were re-scaled according to the intensity of the internal standard (the 30 mer was used for the 10% gel and the 98 mer was used for the 8% gel). Background at each nucleobase position, reflected in the -E control, was subtracted from the glycosylase-treated samples. At each U site the ratio of corrected band intensity in NCPs to that in duplex DNA was used to reveal the ratio of product yield ($P_{\text{NCP}}/P_{\text{dup}}$), and five replicates of the ratio were collected. The standard error (SE) of $P_{\text{NCP}}/P_{\text{dup}}$ was calculated using

$$\text{SE} = \frac{\sigma}{\sqrt{n}}$$  \hspace{1cm} (3.1)

where $\sigma$ is the standard deviation and $n$ is the sample size ($n=5$). A two-tailed student’s $t$ test was performed to obtain the $p$ value at each U site for $P_{\text{variant NCP}}/P_{\text{dup}}$ in comparison to $P_{\text{canonical NCP}}/P_{\text{dup}}$. We considered $p < 0.05$ to be significant.

### 3.3.4 Single turnover kinetics experiments

For each time course, the glycosylase reaction was initiated by manually combining 8 $\mu$L of 40 nM substrate (either duplex DNA or NCPs containing a U at site 110) with equal volume of 1.28 $\mu$M UDG or SMUG1 in reaction buffer (20 mM Tris-HCl [pH 7.6], 50 mM NaCl, 150 mM KCl, 1 mM EDTA, 1 mM DTT, 200 $\mu$g/mL BSA), proceeded at 37 °C for varied amounts of time (5 s - 60 min), and was manually quenched by the addition of 16 $\mu$L of 1 M NaOH. Meanwhile, a negative control sample (-E) was prepared by combining 8 $\mu$L of 40 nM substrate with equal volume of the reaction buffer and incubating at 37 °C for the duration of the longest time point, followed by the addition of 16 $\mu$L of 1 M NaOH. All samples were then heated at 90 °C for 2 min to induce strand breaks at abasic sites. DNA fragments were isolated from proteins
by extraction against PCI. To desalt samples, ethanol precipitation was performed by supplementing the resulting aqueous phase with 600 µL of ethanol and 20 µL of the co-precipitation reagent. Samples were subjected to 8% denaturing PAGE, imaged by phosphorimagery, and quantified by densitometry (Bio-Rad Quantity One). The fraction of product at each time point, $F_P(t)$, was determined by the formula:

$$F_P(t) = \frac{\delta_P(t)}{\delta_S(t) + \delta_P(t)}; \quad (3.2)$$

where $\delta_S(t)$ and $\delta_P(t)$ are the densities of the substrate and product bands, respectively, at time $t$. Considering the pre-existing and incidental substrate damage, $F_P(t)$ was corrected using the following formula:

$$F_{P}^{\text{cor}}(t) = \frac{F_P(t) - F_P(-E)}{1 - F_P(-E)}; \quad (3.3)$$

where $F_{P}^{\text{cor}}(t)$ is the corrected $F_P(t)$ and $F_P(-E)$ is the product fraction observed in the -E sample (generally less than 2%). The mean product fraction from three independent trials was determined at each time point. Based on the averaged data, the time course of product fraction was then fit to a single- or double-exponential model using the following formula to determine the first-order rate constant $k_{\text{obs}}$:

$$F_{P}^{\text{cor}}(t) = F_P(t)|_{t \rightarrow \infty}(1 - e^{-k_{\text{obs}}t}), \quad (3.4)$$

or

$$F_{P}^{\text{cor}}(t) = F_{P1}(t)|_{t \rightarrow \infty}(1 - e^{-k_{1\text{obs}}t}) + F_{P2}(t)|_{t \rightarrow \infty}(1 - e^{-k_{2\text{obs}}t}), \quad (3.5)$$

where $F_P(t)|_{t \rightarrow \infty}$ is the maximum product fraction, using a nonlinear least squares regression (Kaleidagraph). All fittings have $R^2$ value over 0.97.
3.3.5 Hydroxyl radical footprinting

Hydroxyl radical footprinting (HRF) reactions were performed based on previously published methods (33, 34). Briefly, 7.5 µL of each 10 mM Fe(II)-EDTA, 10 mM sodium ascorbate, and 0.12% w/v aqueous hydrogen peroxide were combined and immediately the reagent mixture was added to 5 pmol of duplex DNA or NCPs in 52.5 µL of buffer (10 mM Tris-HCl [pH 7.5], 1 mM EDTA). The reaction proceeded in dark at ambient temperature for 2 min (duplex DNA) or 10 min (NCPs) and was quenched with the addition of 16 µL 50 mM EDTA in 25% v/v glycerol. Duplex DNA was precipitated with the addition of 50 µL of 7.5 M NH₄OAc and 600 µL of ethanol. NCP samples were immediately applied to 7% native PAGE for separating NCPs from duplex DNA that disassociated during the reaction. The band containing NCPs was incised and NCPs were eluted into buffer (0.3 M NaOAc, 1 mM Tris-HCl [pH 8.0], 1 mM EDTA) overnight at 37 °C with gentle shaking. The eluent was concentrated and extracted twice against PCI. DNA fragments in the resulting aqueous phase was purified via ethanol precipitation. Samples were resuspended in 50% v/v formamide, split in half, and half loaded onto a 10% denaturing PAGE gel for resolving bands from nucleobase 11 to the dyad axis and half loaded onto an 8% gel for resolving bands from the dyad axis to nucleobase 140. Gels were imaged by phosphorimagery, and bands were quantified using SAFA software. To categorize the solution accessibility of each nucleobase, we first identified the highest HRF reactivity within a helical turn of nucleosomal DNA. The ratio of band intensity at each nucleobase position within this helical turn was then obtained by dividing the HRF value at a given position by the highest HRF reactivity. Positions with a ratio greater than 0.8, ranging between 0.8-0.2, and below 0.2 were assigned as sites that have high, intermediate, and low solution accessibility, respectively.
3.3.6 Micrococcal nuclease digestion

1 pmol of NCPs was treated with 1, 4 or 20 units of Micrococcal nuclease (MNase) (New England Biolabs) in 20 µL of reaction buffer (50 mM Tris-HCl [pH 7.5], 5 mM CaCl₂, 100 µg/mL BSA) for 5 min at ambient temperature. MNase was inactivated by incubation at 37 °C for 30 min with the addition of 20 µL of the stop solution (50 mM Tris-HCl [pH 7.5], 200 mM EDTA, 0.25% w/v SDS and 0.5 mg/mL proteinase K). Following the incubation, DNA fragments were extracted against PCI and desalted by ethanol precipitation. 8% denaturing PAGE was used to visualize and analyze the cleavage of nucleosomal DNA from MNase digestion.

3.3.7 Analysis of differential accessibilities of macroH2A octasomes and hexasomes to UDG and SMUG1

20 nM macroH2A NCPs containing a U at site 110 and 640 nM UDG or SMUG1 in 16 µL of reaction buffer (20 mM Tris-HCl [pH 7.6], 50 mM NaCl, 150 mM KCl, 1 mM EDTA, 1 mM DTT, 200 µg/mL BSA) were incubated at 37 °C for 60 min, and were immediately loaded on a 7% native PAGE gel for separating octasomes from hexasomes. Bands containing the hexasomes, octasomes, and SMUG1-hexasome complexes were quantified by densitometry, individually incised and eluted into buffer (0.3 M NaOAc, 1 mM Tris-HCl [pH 8.0], 1 mM EDTA). The eluent was concentrated and extracted against PCI. The resulting aqueous phase was supplemented with equal volume of 1 M NaOH and heated at 90 °C for 2 min and was desalted by ethanol precipitation. 8% denaturing PAGE was used to resolve the composition of samples.
3.3.8 Molecular modeling

The 601 macroH2A-containing octasome model was prepared by aligning crystal structures of canonical NCP formed from the Widom 601 sequence and *X. laevis* histones (PDB ID 3lz0) and NCP formed from α-satellite DNA, human macroH2A (histone domain), and mouse H2B, H3 and H4 (PDB ID 1u35) using the molecular graphics and modeling package PyMOL. The alignment was performed by superimposing the histone core of the two crystal structures, since histones from *X. laevis* and mouse share strong homology in protein sequence and structure. To construct the hexasome model the macroH2A-H2B dimer close to the 3’-end of the I strand was removed. The model for 601 H2A.Z-containing octasome was constructed similarly using the crystal structure of NCP formed from α-satellite DNA, mouse H2A.Z, and *X. laevis* H2B, H3 and H4 (PDB ID 1f66).

3.4 Results

3.4.1 Characterization of NCPs with global C to U substitution

To investigate the global profile of U removal from canonical NCPs or those containing histone H2A variants, we prepared NCPs using “Widom 601” DNA (35). This positioning sequence possesses high affinity for and binds the histone octamer in a single orientation (36). U lesions were globally incorporated into the “I strand” of Widom 601 DNA using chemical synthesis techniques to create U:G bp, mimicking the natural consequence of cytosine deamination. The global incorporation of U was achieved using a mixture of C/U building blocks during DNA synthesis with the molar ratio determined by a Poisson distribution, ensuring that 95% of the DNA contained
at most one U per strand. Incubating the single-stranded DNA with UDG confirmed that the strategy provides a population of DNA where collectively each C is replaced with U (Figure 3.1).

NCPs were reconstituted with DNA containing global U lesions and histone octamers via salt gradient dialysis. Since homotypic H2A.Z NCPs have been determined to be enriched at active genes in *Drosophila* (37) and to exist exclusively (rather than heterotypic NCPs) in *Arabidopsis thaliana* (38), we reconstituted homotypic H2A.Z and macroH2A containing NCPs, in which both copies of canonical H2A were replaced by the variant protein. Similar to other studies that examined the influence of macroH2A incorporation on NCPs we used the histone domain of the variant (39–42). Formation of NCPs and their purity was evaluated by native PAGE (Figure 3.2). While canonical and H2A.Z NCPs migrate as a single species, macroH2A NCPs migrate as two species. A similar observation was reported previously for macroH2A NCPs (40).

To determine if these two species arise from different translational positioning of the DNA, we performed a heat-shifting assay where NCPs were incubated at 37 °C or 55 °C (Figure 3.3) (43). No heat-induced redistribution of the NCP species was observed, indicating that the two species are thermodynamically stable under the experimental conditions with no interchange between the two species. We therefore hypothesized that macroH2A NCPs exist as two distinct populations.

Hydroxyl radical footprinting (HRF) was then utilized to establish the solution accessibility of U lesions in NCPs. Nucleobases that face outward are highly susceptible to hydroxyl radicals, resulting in strand cleavage, while nucleobases that face in toward the histone protein core are sequestered and protected from cleavage. The expected oscillatory pattern of cleavage is observed in the HRF profiles for all NCPs (Figure 3.4A and B). To categorize the solution accessibility of each
Figure 3.1 Confirmation of global U incorporation into Widom 601 DNA. Single-stranded DNA containing global C to U substitution was incubated with UDG followed by NaOH-catalyzed strand break at abasic sites. Lane 1 shows the Maxam-Gilbert sequencing ladder (A+G). Lanes 2 and 3 are two different DNA syntheses, and lane 4 shows a no treatment control of the single strand. Anticipated U are indicated on the right.
Figure 3.2 Representative native PAGE of reconstituted NCPs. Single- and double-stranded DNA controls are shown in lanes 1 and 2. Lanes 3 to 5 indicate the migration of canonical, H2A.Z and macroH2A NCPs, respectively.

Figure 3.3 Heat-shifting assay on canonical and H2A variant NCPs. Canonical, H2A.Z and macroH2A NCPs were incubated separately at 37 °C or 55 °C for one hour and immediately subjected to native PAGE at 4 °C. Non-treated NCPs (lanes 1, 4 and 7) and a double-stranded DNA control (lane 10) are included.
nucleobase, we first identified the highest HRF reactivity within a helical turn of nucleosomal DNA. The ratio of band intensity at each nucleobase position within this helical turn was then obtained by dividing the HRF value at a given position by the highest HRF reactivity. Positions with a ratio greater than 0.8, ranging between 0.8-0.2, and below 0.2 were assigned as sites that have high, intermediate, and low solution accessibility, respectively. Notably, U lesions are located at diverse translational positions and with varied levels of solution accessibility (Figure 3.4C), allowing us to evaluate U removal from NCPs on a global level.

3.4.2 Enhanced excision of U from NCPs containing H2A variants

We next investigated U removal by UDG and SMUG1 in the context of NCPs. We used single-turnover conditions where $[\text{glycosylase}] + [\text{substrate}]$. Since UDG and SMUG1 are able to excise U from duplex DNA (46, 47), we included this substrate as a positive control for enzymatic activity. Figure 3.5 and 3.6 show the denaturing PAGE gel visualization of U excision in duplex DNA and NCP substrates, by UDG and SMUG1, respectively.

At each U site the ratio of product yield (after 60 min-reaction) in NCPs to that in duplex DNA was plotted versus nucleobase position. A ratio of 1 indicates that the lesion can be excised as efficiently in NCPs as in duplex DNA. We observed that UDG activity in canonical NCPs is largely dominated by solution accessibility of U (Figure 3.7A, red bars). Multiple sites with high solution accessibility, as determined by HRF, exhibit high UDG activity where ratios exceed 0.6 (sites 22, 23, 44, 55, 63, 94, 95, 106 and 115). In a similar manner, sites that face toward the histone protein core have low solution accessibility, and as expected, display minimal UDG
Figure 3.4 Solution accessibility of nucleosomal DNA. (A) Hydroxyl radical footprinting (HRF) of DNA in canonical and H2A variant NCPs visualized by denaturing PAGE. Lanes 1 to 3 are HRF results for canonical (C NCP), H2A.Z (Z NCP) and macroH2A NCP (M NCP), respectively. Lane 4 shows the Maxam-Gilbert sequencing ladder (A+G). (B) Band intensity at each nucleobase position on the denaturing PAGE gel in panel A was quantified, normalized, and plotted versus nucleobase position, illustrating the varying solution accessibility along the sequence in NCP samples (canonical NCP: red; H2A.Z NCP: blue; macroH2A NCP: dark green). Representative sites of high, intermediate (inter.) and low solution accessibility are indicated. (C) Model of a canonical NCP with U globally incorporated at diverse translational positions and with varied levels of solution accessibility. The model was constructed by merging the crystal structure of an NCP containing Widom 601 DNA with a canonical histone octamer containing histone tails (PDB ID 3lz0 (44) and 1kx5 (45), respectively). U sites are highlighted according to their solution accessibility that are determined by HRF (high: blue; intermediate: purple; low: red). The dyad axis is indicated by a dashed line.
Figure 3.5 UDG activity on Widom 601 duplex DNA and NCPs containing global C to U substitution. 25 nM duplex DNA or NCPs were incubated with 500 nM UDG for 10 min or 60 min. NaOH-catalyzed strand breaks at abasic sites after UDG reactions reveal the pattern of U excision in varied substrates. (A) 10% denaturing PAGE visualization of U excision. In each substrate data set, lanes are as follows: no treatment control (NT lane) in which no UDG or NaOH treatment was performed; negative control (-E lane), treated with NaOH only, indicating any pre-existing and incidental damage occurred before the experiments or during sample workup; U excision after 10 min-UDG incubation; U excision after 60 min-UDG incubation. The internal standards are indicated with an asterisk on the gel. Bands from nucleobase 9 to the dyad axis are quantified by SAFA. (B) same analysis of U excision as in panel A on a 8% denaturing PAGE gel. Bands from the dyad axis to nucleobase 129 are quantified.
Figure 3.6 SMUG1 activity on Widom 601 duplex DNA and NCPs containing global C to U substitution. 25 nM duplex DNA or NCP substrates were incubated with 500 nM SMUG1 for 60 min. NaOH-catalyzed strand breaks at abasic sites after SMUG1 reactions reveal the pattern of U excision in varied substrates. (A) 10% denaturing PAGE visualization of U excision. Each substrate data set contains a SMUG1-incubated sample and a negative control (-E lane), which is treated with NaOH only and thus indicates any pre-existing and incidental damage occurred before the experiments or during sample workup. The internal standards are indicated with an asterisk on the gel. Bands from nucleobase 9 to the dyad axis are quantified by SAFA. (B) same analysis of U excision as in panel A on a 8% denaturing PAGE gel. Bands from the dyad axis to nucleobase 129 are quantified.
activity with ratios below 0.2 (sites 16, 17, 27, 46, 61, 69, 81, 91, 92, 110 and 122). The region flanking the dyad axis, spanning from site 70 to site 92, is a notable exception to solution accessibility dictating glycosylase activity; excision of U by UDG is significantly inhibited regardless of solution accessibility in this region.

In the macroH2A-containing NCPs, the correlation between solution accessibility of U and UDG activity is generally retained, with some exceptions. In Figure 3.7B, UDG activity in variant NCPs is compared to that in canonical NCPs; a ratio of or greater than 1 indicates comparable or enhanced U excision in variant NCPs, respectively. While most highly solution-accessible U sites still exhibit high UDG reactivity, 1.5- to 10-times more product is observed at all sites with low solution accessibility and at 11/18 sites with intermediate solution accessibility compared to canonical NCPs (Figure 3.7B, green dots). Moreover, whereas U removal is still inhibited in the dyad region relative to the rest of the sequence, an overall increase in UDG activity is worth noting, especially at sites 73 and 77 with ratios reaching 0.4 (Figure 3.7A, green bars). Quite intriguingly, starting from site 108 and to the last examined lesion site 129, there is efficient U excision at all lesion sites in this region, with a remarkable increase in UDG activity at sites 109, 110 and 122, relative to canonical NCPs.

Similar to the influence exerted by macroH2A, the incorporation of H2A.Z gives rise to 2 to 5-times higher UDG activity at most lesion sites with low solution accessibility and half of the sites with intermediate solution accessibility relative to canonical NCPs (sites 25, 27, 35, 46, 65, 109, and 118) (Figure 3.7B, blue dots). Additionally, a small but statistically greater amount of U removal is observed at most sites in the dyad region of H2A.Z NCPs (Figure 3.7A, blue bars). Nevertheless, the striking enhancement in UDG activity observed for macroH2A at sites 109, 110, and 122 is not observed with H2A.Z.
Figure 3.7 Excision of U from NCPs containing global C to U substitution after 60 min-incubation with glycosylases. (A) U excision initiated by UDG. At each U site the ratio of product yield in NCPs to that in duplex DNA is plotted versus nucleobase position (canonical NCP: red; H2A.Z NCP: blue; macroH2A NCP: green). The ratio of 1 is indicated as a dashed line and reflects a case where U excision from NCPs is as efficient as in duplex DNA. Position of the dyad axis is indicated by an arrow. (B) U excision initiated by UDG. U sites are re-categorized based on their solution accessibility. At each U site the ratio of product yield in H2A.Z NCPs (blue dots) or in macroH2A NCPs (green dots) to that in canonical NCPs is plotted. The ratio over 1 indicates a greater level of U excision from variant NCPs than canonical NCPs. (C) U excision initiated by SMUG1. At each U site the ratio of product yield in NCPs to that in duplex DNA is plotted versus nucleobase position (canonical NCP: red; H2A.Z NCP: blue; macroH2A NCP: green). Error bars represent the standard error (n=5).
Unlike UDG, SMUG1 activity is substantially suppressed along the entire sequence in canonical NCPs (Figure 3.7C, red bars); efficient U excision by SMUG1 is only detected at site 11, which is moderately exposed to solution and close to the DNA entry/exit region, and at site 95, which is highly solution-accessible. Of particular interest, within macroH2A NCPs restoration of glycosylase activity is manifested for SMUG1 as well in the region from site 106 to 129, revealed by the high ratios of product yield (ranging between 0.45 and 0.85) at all U sites (Figure 3.7C, green bars). The incorporation of H2A.Z, on the other hand, has minimal influence on U excision and nearly the same level of repair is observed for canonical and H2A.Z NCPs (Figure 3.7C, blue bars).

### 3.4.3 Kinetics of U excision by UDG and SMUG1 from H2A variant NCPs

To further understand the effect of H2A variants on U removal from NCPs, we performed kinetic experiments on substrates containing a single C to U substitution. Site 110 was selected for the lesion position since it is a site where both UDG and SMUG1 exhibit a drastic difference in reactivity in canonical versus H2A variant NCPs. Moreover, site 110 is in close proximity to the L1-L1' interface of the two H2A-H2B dimers. Distinct structural properties of the L1-L1' interface has been proposed in macroH2A and H2A.Z NCPs to confer specialized biological functions (16). The experiment was performed under single-turnover conditions, and the observed rate \( k_{obs} \) reflects the slowest step up to and including cleavage of the glycosidic bond (48). Excision of U by the two glycosylases were visualized by denaturing PAGE gels (Figure 3.8A and B).

Consistent with earlier reports, UDG excises U from 145 bp duplex DNA
with a $k_{\text{obs}}$ of 21 min$^{-1}$ (Table 3.1) (24, 47). In contrast, UDG activity is diminished in canonical and H2A.Z NCPs, as evidenced by a maximum of 8% and 16% product yield, respectively. As site 110 exhibits low solution accessibility, this decrease in UDG reactivity in NCP is not surprising and is likely caused by steric obstruction by the protein core. In macroH2A NCPs, however, remarkable restoration of UDG activity is observed with the product yield increasing to 66%. It is notable that product yields observed in this site-specific kinetics study agree with those obtained for site 110 using the global U-containing NCPs (Figure 3.7A), validating the feasibility of the approach in which U removal in NCPs can be assessed both globally and quantitatively.

Interestingly, UDG exhibits two kinetic phases for all NCP substrates. While product formation for the slow phase is 150- to 500-times slower than in duplex DNA, rates of the fast phase are comparable to $k_{\text{obs}}$ for duplex DNA, suggesting a population that is nearly as accessible as duplex. Nevertheless, since the fast phase accounts for a very small fraction of product in canonical and H2A.Z NCPs (~3%), it is possible that the fast phase results from a small amount of duplex DNA rather than from authentic NCP substrates. However, this cannot be applicable for macroH2A NCPs, as a majority of the product (71%) in macroH2A NCPs is obtained in the fast phase, indicating a distinct population that is amenable to processing by UDG.

Compared to UDG, SMUG1 excises U from duplex DNA at a much slower rate ($k_{\text{obs}} = 0.55$ min$^{-1}$) and is significantly impaired with less than 5% product formation observed for canonical and H2A.Z NCPs. With macroH2A NCPs, a product yield of nearly 60% is observed, most of which is contributed by the fast phase ($k_{\text{obs}} = 0.38$ min$^{-1}$).
Figure 3.8 Single-turnover kinetics time courses of UDG and SMUG1 acting on duplex DNA and NCPs containing a single C to U substitution at site 110. Representative 8% denaturing PAGE gels showing the conversion of substrate (Sub) to product (Pdt) during kinetic assay for (A) UDG and (B) SMUG1. (C) Reaction time courses for UDG and SMUG1 acting on duplex DNA (black) and NCPs (canonical NCP: red; H2A.Z NCP: blue; macroH2A NCP: green). Lines are the best fit to the appropriate single- or double-exponential growth models. Error bars represent the standard error (n=3).
Table 3.1. Kinetic parameters of U excision at site 110 from duplex DNA and NCPs.

<table>
<thead>
<tr>
<th>Glycosylase</th>
<th>Substrate</th>
<th>$k_{obs}/\text{min} \ (%\text{product})^a$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Fast phase</td>
</tr>
<tr>
<td>UDG</td>
<td>duplex DNA</td>
<td>$21.0 \pm 0.4 \ (99%)$</td>
</tr>
<tr>
<td></td>
<td>canonical NCP</td>
<td>$6.5 \pm 2.4 \ (3.6%)$</td>
</tr>
<tr>
<td></td>
<td>H2A.Z NCP</td>
<td>$16 \pm 4 \ (3.4%)$</td>
</tr>
<tr>
<td></td>
<td>macroH2A NCP</td>
<td>$6.7 \pm 0.5 \ (47%)$</td>
</tr>
<tr>
<td>SMUG1</td>
<td>duplex DNA</td>
<td>$0.55 \pm 0.02 \ (97%)$</td>
</tr>
<tr>
<td></td>
<td>canonical NCP</td>
<td>$1.4 \pm 0.5 \ (2.0%)$</td>
</tr>
<tr>
<td></td>
<td>H2A.Z NCP</td>
<td>$0.23 \pm 0.05 \ (2.8%)$</td>
</tr>
<tr>
<td></td>
<td>macroH2A NCP</td>
<td>$0.38 \pm 0.07 \ (38%)$</td>
</tr>
</tbody>
</table>

$^a$Error represents standard deviation from fitting by weighted nonlinear least-squares regression.

3.4.4 MacroH2A NCPs are a mixture of octasomes and hexasomes

We next addressed the question of why UDG and SMUG1 activities are restored in a defined region of macroH2A NCPs, sites 108-129. Since two distinct species were observed by native PAGE, we performed HRF reactions on these two species separately to investigate their architecture. Visualization and quantification of strand cleave after HRF reaction are shown in Figure 3.9 and Figure 3.10, respectively. For the faster-migrating species, which co-migrates with canonical NCPs, the oscillatory pattern of cleavage is observed along the entire length of DNA (Figure 3.10, gray curve). In contrast, starting from site 115 to site 140, the slower-migrating species exhibits high susceptibility towards hydroxyl radicals irrespective of the rotational orientation relative to histones (Figure 3.10, green curve), indicating much weaker interaction or loss of contact between the DNA and the protein core. Importantly,
this region is where we observed restoration of UDG and SMUG1 activity. The slower-migrating species possesses a less compact conformation, in which only ~115 bp of DNA is tightly associated with the histone core. Notably, this observation is consistent with earlier studies of a hexasome, which lacks one copy of H2A-H2B dimer and wraps only 112 bp of DNA (49, 50). Furthermore, we note that the macroH2A-H2B dimer is preferentially depleted in the area that is near the 3′ end of the I strand in the hexasomes, since this is the only region where we observed overall increased susceptibility of nucleosomal DNA towards hydroxyl radicals.

To confirm a portion of macroH2A NCPs forms hexasomes, we treated the NCPs with micrococcal nuclease (MNase), which preferentially digests DNA that is not closely associated with the histone core. For canonical NCPs, two distinct DNA fragments, a 128 mer and a 132 mer, were observed at high MNase concentration (Figure 3.11), indicating that 13-17 bp of DNA were digested at the 3′ end of the I strand, likely due to the transient unwrapping of DNA in the entry-exit region. Additionally, cleavage at sites 95, 106, and 126 was observed as these sites are highly solution-accessible. In comparison, while macroH2A NCPs are similarly accessible to MNase at sites 95, 106, 126, 128 and 132, the cleavage at sites 102, 103, and 110 is unique to macroH2A NCPs. Notably, the 110 mer is detected for macroH2A NCPs even at low MNase concentration, indicating a weak interaction between the histone core and the 35 bp of DNA at the 3′ end of the I strand. Presence of the 102 mer and 103 mer in the digested macroH2A NCPs may result from the transient interactions of nucleosomal DNA close to the DNA entry-exit region in hexasomes. These findings, in accord with the HRF results, strongly suggest that the macroH2A NCPs in our experiments contain two species: octasomes and hexasomes.
Figure 3.9 Hydroxyl radical footprinting of nucleosomal DNA. Denaturing PAGE visualization of NCP HRF results. Bands are quantified by SAFA, with nucleobase 7 to 70 resolved by a 12% gel (left) and nucleobase 71 to 140 resolved by a 8% gel (right). On each gel lanes are as follows: Maxam-Gilbert sequencing ladder (A+G) (L); canonical NCP (C); H2A.Z NCP (Z); macroH2A hexasome (M₉); macroH2A octasome (M₀).
Figure 3.10 Hydroxyl radical footprinting of DNA in canonical and H2A variant NCPs. Band intensity at each nucleobase position on the denaturing PAGE gel (Figure 3.9) was quantified and normalized, illustrating the varying solution accessibility along the sequence in each NCP sample (canonical NCP: red; H2A.Z NCP: blue; macroH2A hexasome: green; macroH2A octasome: gray). Position of the dyad axis is indicated by an arrow. Representative sites with high, intermediate (inter.) and low solution accessibility are shown.

Figure 3.11 Micrococcal nuclease (MNase) digestion of canonical and macroH2A NCPs. Canonical and macroH2A NCPs were treated with 1 unit (lanes 2 and 6), 4 units (lanes 3 and 7) and 20 units (lanes 4 and 8) of MNase. The negative controls in which no MNase was added are shown in lanes 1 and 5. The digestion results were resolved by denaturing PAGE. Lane 9 is a size ladder created by performing the Maxam-Gilbert sequencing reaction (A+G) on the Widom 601 J strand. Cleavage sites that are observed exclusively in macroH2A NCPs are indicated by arrows.
3.4.5 MacroH2A octasomes and hexasomes both exhibit increased accessibility to UDG and SMUG1

The activities of UDG and SMUG1 on individual macroH2A hexasome and octasome species were next examined. After incubation with UDG or SMUG1, macroH2A NCPs containing a U at site 110 were subjected to native PAGE for separating octasomes from hexasomes (Figure 3.12A). Given the strong product inhibition of SMUG1, a new band appeared concomitantly with disappearance of the hexasome band, likely the SMUG1-bound hexasome complex. The level of U excision from each species was then analyzed by denaturing PAGE (Figure 3.12B and 3.12C). We found that both the hexasome and the octasome species of macroH2A NCPs exhibit increased accessibility to UDG and SMUG1 relative to canonical NCPs. Moreover, while 40% and 20% product formation are observed in macroH2A octasomes with UDG and SMUG1, respectively, 80% product formation is observed in hexasomes with both glycosylases, revealing that the hexasomes are more amenable to processing by these glycosylases than the octasomes.

3.4.6 Mapping of U sites in H2A.Z and macroH2A NCP models

In order to investigate the impact of structural alterations in macroH2A and H2A.Z NCPs on U excision, we categorized U sites depending on the degree of increase in UDG or SMUG1 activity (more than 30% increase, 30 to 10%, less than 10%) and mapped them onto a macroH2A or H2A.Z NCP model. In the macroH2A octasome model, we note that U sites with the most dramatic increase in UDG activity cluster in three regions (Figure 3.13A). One region is near the docking domain at the C-terminus of macroH2A, and the other two regions are near the N-terminus of macroH2A and
Figure 3.12 Differential accessibilities of U at site 110 in canonical and macroH2A NCP species to UDG and SMUG1. (A) Native PAGE analysis following the incubation of macroH2A NCPs with UDG and SMUG1. Lane 1 is a duplex DNA control. Lane 2 shows macroH2A NCPs in the absence of a glycosylase. UDG and SMUG1 treated macroH2A NCPs are shown in lanes 4 and 5, respectively; boxed bands represent the hexasomes (hex), octasomes (oct), and SMUG1-hexasome complex species (comp), and were individually excised and eluted for analysis in panel B. (B) Denaturing PAGE analysis of U excision in macroH2A hexasomes and octasomes. Lane 1 is a size marker containing a 145 mer (substrate) and 110 mer (product). Lanes 2 to 6 show the composition of each resulting eluent from the excised bands in panel A. (C) Quantitation of the denaturing PAGE gel in panel B. Fraction product with UDG or SMUG1 is shown as black and gray bars, respectively. Data for macroH2A hexasomes with SMUG1 is the weighted mean of fraction product of eluent in lanes 4 and 6 in panel B. Data for canonical NCPs is the fraction product at 60-min time point in Figure ???. Error bars represent the standard error (n=3).
the L1-L1’ interface of the two macroH2A-H2B dimers. A similar clustering pattern of U sites was observed for H2A.Z octasomes (Figure 3.14).

Since macroH2A NCPs are a mixture of octasomes and hexasomes, we also created a macroH2A hexasome model by removing the macroH2A-H2B dimer that associates with nucleosomal DNA at the 3’ end of I strand from the octasome (Figure 3.13B). Earlier work demonstrated that the theoretical small-angle X-ray scattering (SAXS) curves of a canonical hexasome model constructed in this way are in good agreement with the experimental SAXS curves of the hexasomes (49). It is noticeable that in the macroH2A hexasome model, the region of DNA from ~site 108 to the DNA terminus, where U sites with the largest increase in excision by UDG and SMUG1 are clustered, has minimal association with the histone core.

3.5 Discussion

In this work, we utilized a quantitative platform for evaluating the global profile of UDG and SMUG1 repair in NCPs. In agreement with earlier work, we found that U excision by UDG from canonical NCPs is correlated with solution accessibility of U (28, 52, 53), but substantially diminished in the dyad region (29, 54). Indeed, a similar pattern of suppressed glycosylase activity in the dyad region has been reported for removal of 8-oxo-7,8-dihydroguanine (8-oxoG) (24) and 7-methyl-guanine (55) from NCPs via the BER pathway. This result is consistent with recent observations in cells of accumulation of oxidative and alkylation damage and high mutation frequency around the dyad axis region (55, 56). Such inhibition of enzymatic activity, observed with restriction enzymes as well (57, 58), may be due to the unusual helical periodicity of nucleosomal DNA in the dyad region (34, 59), and the extremely low frequency of exposing nucleobases via transient unwrapping near the dyad axis (60).
Figure 3.13 Model of a macroH2A (A) octasome and (B) hexasome with global C to U substitution. U sites are colored according to the level of increase in UDG or SMUG1 activity in macroH2A NCPs relative to canonical NCPs: 0 - 10% (blue), 10 - 30% (orange), and 30% and more (pink). All orange and pink sites were determined to have a \( p \) value less than 0.05 in comparison to canonical NCPs. MacroH2A is shown in green, and other histones in gray. The 5' and 3' ends of U-containing strand (I strand) are indicated. In the zoomed in views, canonical H2A (yellow) and macroH2A (green) are superimposed (other histones and the J strand are not included for simplification). The docking domain, L1-L1' interface and the N-terminus of H2A/macroH2A are indicated.
Figure 3.14 Model of an H2A.Z NCP with global C to U substitution. The model was constructed by merging an NCP structure containing the Widom 601 sequence (PDB ID 3lz0 (44)) and another NCP structure containing the H2A.Z octamer (PDB ID 1f66 (51)). U sites are colored according to the level of increase in UDG activity in H2A.Z NCPs relative to canonical NCPs: 0 - 10% (blue), 10 - 30% (orange), and 30% and more (pink). All orange and pink sites were determined to have a \( p \) value less than 0.05 in comparison to canonical NCPs. H2A.Z is shown in cyan, and other histones in gray. In the zoomed in view, canonical H2A (yellow) and H2A.Z (cyan) are superimposed (other histones and the J strand are not included for simplification). The docking domain, L1-L1' interface and the N-terminus of H2A/H2A.Z are indicated.
Different observations, though, have been made in other global assessments of U excision from NCPs. We attribute these differences to the varied experimental conditions and the complexity of NCP substrates, which derives from the sequence context, base pair composition and source of histone proteins that differ in the level of PTMs. Ye and co-workers reported that, for U:A bp in Widom 601 DNA, it is predominately local DNA structure that determines U excision under multiple-turnover conditions, with rotational orientation playing a minor role at several sites (26). Also using U:A bp, Nilsen et al. observed excision of U to be only marginally reduced in NCPs assembled with 5S rDNA and chicken erythrocyte histones relative to duplex DNA, irrespective of rotational orientation (61).

Compared with UDG, we observed that repair initiated by SMUG1 in canonical NCPs is mostly abolished. A co-crystal structure of SMUG1 and duplex DNA shows an invasive interaction that may not be possible in NCPs (62). Structural constraints of nucleosomal DNA and obstruction imposed by histone proteins add other layers of difficulty for SMUG1 to process U lesions.

Due to the hindrance of canonical histones to DNA repair machinery, histone variants provide a strategy to facilitate DNA repair and maintain genomic stability (63). For instance, there is extensive evidence for the involvement of H2A.X, and the phosphorylated version γ-H2A.X, in response to DSB (64, 65). Furthermore, the deposition of H2A.Z has been reported for remodeling chromatin architecture and recruiting machinery for DSB repair (66), as well as for NER (19). MacroH2A is also found at DSB sites and cells lacking macroH2A exhibit increased radiosensitivity (23). However, little is known about the potential roles of histone variants in BER. To our knowledge, one of the only two reports of histone variants in BER showed increased sensitivity of cells to methyl methanesulfonate treatment after macroH2A depletion (23), implying macroH2A functioning in the repair of methylated nucleobases via
BER. The other one used H2A.Bbd in studies of human oxoguanine glycosylase 1 (OGG1). The authors observed minimal excision of 8-oxoG that was positioned near the dyad axis in both canonical and H2A.Bbd NCPs (67).

Upon replacing canonical H2A with macroH2A or H2A.Z in our experiments, we observed globally enhanced reactivity of UDG in NCPs, particularly at sites with reduced solution accessibility. Of interest, in both macroH2A and H2A.Z NCPs, sites with the largest increase in UDG efficiency are clustered and are near the docking domain, the N-terminus, and the L1 loops of the two variants. As the docking domain of the H2A variants exhibit extensive interactions with αN and α2 of H3 (Figure 3.13A), which associate with DNA at the DNA entry/exit region and at the dyad axis respectively, it is likely that sequence divergence in the docking domain of H2A variants strategically modulates DNA dynamics and accessibility in these regions (16). This effect, indeed, is also manifested in the enhanced U excision by SMUG1 in H2A.Z NCPs at the DNA entry/exit region (sites 9-11). The docking domain of H2A.Z adopts fewer hydrogen bonds with H3, leading to the destabilization of local structure. The L1 loops, being the only interface between the two H2A-H2B dimers, has been suggested to play a critical role in regulating DNA-histone binding (68) and the dynamics and energetics of macroH2A and H2A.Z NCPs (39, 51). Consequently, structural rearrangement of the L1 loops in variant NCPs could contribute to the clustering of U sites with the largest increase in UDG activity around the L1-L1’ interface. Based on a recent study of H2A variants in Arabidopsis thaliana (38), and together with previous molecular simulations (68), the impact of the distinct docking domain and L1 loops of H2A variants can alter not only the structure and dynamics of individual NCP, but also inter-nucleosomal interactions. Therefore, incorporating H2A variants may be profoundly influential on the accessibility of chromatin fibers. It is also of note that due to the strong association between the Widom 601 sequence
and the histone core, our observations may reflect an underestimate of the extent to which U can be repaired in other sequences; the ability of macroH2A and H2A.Z to facilitate BER may be even more dramatic than the significant effect we observed.

Another intriguing finding is the restoration of SMUG1 activity at all lesion sites in a defined region of macroH2A NCPs, where high UDG efficiency is also observed. We attribute this recovery in glycosylase activity to weaker or complete loss of DNA-histone interactions in this region, revealed by both HRF and MNase treatment. Several scenarios can result in weaker DNA-histone interplay and a less compact NCP structure, one of which is the formation of end-positioned NCPs that includes macroH2A (69). However, formation of the end-positioned NCPs with Widom 601 sequence requires the assistance of chromatin remodelers (69). Alternatively, as there has been growing evidence for preferential unwrapping of DNA from the 3'-end of the I strand in Widom 601 canonical NCP, as discussed in §1.6.3 (70–72), asymmetric unwrapping can be another possibility. Nevertheless, rates of lesion site exposure induced by unwrapping alone are too low to account for observed rates of U excision at site 110 (60, 73). Therefore, it is mostly likely that the formation of macroH2A hexasomes contributes to the restoration of glycosylase activity in the region 35-40 bp from the 3'-end of the I strand, and this contribution could be amplified by synergistic asymmetric unwrapping occurring within the hexasomes.

In fact, it has been reported that hexasomes can concomitantly form with octasomes during canonical NCP reconstitution (74), and a macroH2A hexasome species was proposed to have been observed (40). Intriguingly, we note that the macroH2A hexasomes appear to lack exclusively one copy of the macroH2A-H2B dimer at the 3'-end of the I strand, where the asymmetrical unwrapping of nucleosomal DNA in Widom 601 canonical NCPs is thermodynamically favored (70, 71, 75).
Previous biochemical experiments indicate that macroH2A octamers are less sensitive to the decreased salt concentration and undergo a peculiar dissociation pathway during reconstitution (40). The distinct property of macroH2A octamers makes it possible that an incompletely dissociated species, the hexamers, are deposited onto DNA, and the more flexible 5'-end of the I strand preferentially binds with the hexamers. Alternatively, as the formation of canonical hexasomes in vitro is reported to be guided by asymmetric unwrapping (72), macroH2A octasomes may have a higher probability than canonical octasomes to release a macroH2A-H2B dimer at the 3'-end of the I strand during reconstitution. Given its distinct structure and dynamics, canonical hexasomes have emerged as an intermediate for modulating chromatin architecture and DNA accessibility during replication, transcription, and DNA repair (76–81). Notably, the finding of histone chaperone FACT, which promotes the RNA polymerase II-dependent H2A-H2B dimer eviction (82), to co-localize with OGG1 at damage sites may imply the involvement of the hexasomes in BER (54).

In our kinetic experiments, we observed biphasic kinetics in most NCPs with UDG and SMUG1. The biphasic kinetics have been reported previously for other glycosylases on NCPs (26, 27, 47). Interestingly, \( k_{\text{obs}} \) of the fast phase observed with NCPs is comparable to that of duplex DNA, indicating a substrate population that is nearly as accessible as duplex. Although the small amount of U excision (~3%) in the fast phase may derive from duplex DNA, native PAGE demonstrates that NCP samples remain intact after UDG and SMUG1 reactions (Figure 3.12A). Furthermore, single molecule FRET studies revealed the existence of an open state of NCPs, populated 0.2-3% under physiological conditions (83), that may possess more accessible DNA and explain the fast phase observed with canonical and H2A.Z NCPs. In contrast, ~50% and 40% of U excision were obtained in the fast phase with UDG and SMUG1, respectively, in macroH2A NCPs. We attribute this increase mostly to
the dramatically enhanced DNA accessibility caused by the macroH2A hexasomes.

U excision in the slow phase of NCPs, on the other hand, is \(\sim 200\)-times and \(\sim 10\)-times slower with UDG and SMUG1, respectively, than in duplex DNA, suggesting a population of NCPs that is comparatively refractory to processing by the glycosylases. We note that while rates of the fast phase in NCPs are dramatically different for UDG and SMUG1, rates of the slow phase are relatively similar, perhaps suggesting that the NCP population accounting for the slow phase undergoes a rate-limiting structural alteration(s) to expose the lesion site or adopt a conformation permissive for repair. In comparison, NCPs in the fast phase may have the same rate-limiting step as duplex DNA since their rates are comparable under single-turnover conditions, and the rate-limiting step is determined by mechanisms unique to UDG and SMUG1, rather than by NCP conformational change(s).

Although exchanging canonical H2A with macroH2A or H2A.Z can modulate the structure and dynamics of NCPs and enhance the efficiency of BER, coupling the incorporation of histone variants with other cellular factors, such as chromatin remodelers and/or histone PTMs, could provide additional and more versatile strategies for regulating DNA accessibility and maintaining genomic integrity. Indeed, H2A.Z sumoylation has been suggested to play a role in DSB repair (84), and acetylated H3.3/H2A.Z hybrid nucleosomes are known to be enriched at the transcription start sites of active genes (85, 86). Moreover, the C-terminal macro domain of macroH2A is capable of interacting with histone deactylases (39) and binding nicotinamide adenine dinucleotide metabolites (87, 88), implying its diverse role in biological events. Taken together, given the complexity of chromatin, elaborate mechanisms have evolved to balance DNA accessibility and genome stability for a multitude of biological processes, such as replication, transcription, and repair. Our study demonstrates the enhancement of U excision in NCPs containing the H2A variants H2A.Z and macroH2A.
Further studies with additional cellular factors will reveal the extent to which they modulate the influence of histone variants on BER.

### 3.6 Supporting information

**DNA sequences used in the present work (35)**

**LS\textsubscript{U\_global}**

5’- ATC AGA ATC CCG GTG CCG AGG CCG CTC AAT TGG TCG TAG ACA GCT CTA GCA CCG CTT AAA CGC ACG TAC GCG CTG TCC CCC GCG TTT TAA CCG CCA AGG GGA TTA CTC CCT AGT CTC CAG GCA CGT GTC AGA TAT ATA CAT CGA T -3’ (U was incorporated randomly but unbiasedly at each C site during synthesis. Based on the Poisson distribution, synthesized DNA contained at most one U per strand.)

**LS\textsubscript{U\_110}**

5’- ATC AGA ATC CCG GTG CCG AGG CCG CTC AAT TGG TCG TAG ACA GCT CTA GCA CCG CTT AAA CGC ACG TAC GCG CTG TCC CCC GCG TTT TAA CCG CCA AGG GGA TTA CTC CCT AGT CTC CAG GCA CGT GTC AGA TAT ATA CAT CGA T -3’

**LS\textsubscript{U\_110.1}**

5’- ATC AGA ATC CCG GTG CCG AGG CCG CTC AAT TGG TC -3’
**LS\textsubscript{U\textsubscript{110.2}}**

5'- GTA GAC AGC TCT AGC ACC GCT TAA ACG CAC GTA CGC GCT GTC CCC -3'

**LS\textsubscript{U\textsubscript{110.3}}**

5'- CGC GTT TTA ACC GCC AAG GGG ATT ACT CCU TAG TCT CCA GGC ACG TGT CAG ATA TAT ACA TCG AT -3'

**LS\textsubscript{U\textsubscript{110.s12}}**

5'- GTG CTA GAG CTG TCT ACG ACC AAT TGA GCG GCC -3'

**LS\textsubscript{U\textsubscript{110.s23}}**

5'- GGC GGT TAA AAC GCG GGG GAC AGC GC -3'

**CS**

5'- ATC GAT GTA TAT ATC TGA CAC GTG CCT GGA GAC TAG GGA GTA ATC CCC TTG GCG GTT AAA ACG CGG GGG ACA GCG CGT ACG TGC GTT TAA GCG GTG CTA GAG CTG TCT ACG ACC AAT TGA GCG GCC TCG GCA CCG GGA TTC TGA T -3'

**CS.1**

5'- ATC GAT GTA TAT ATC TGA CAC GTG CCT GGA GAC TAG GGA GTA ATC -3'
CS.2

5'- CCC TTG GCG GTT AAA ACG CGG GGG ACA GCG CGT ACA TGC GTT
TAA GCG GTG CTA GAG CTG TCT AC -3'

CS.3

5'- GAC CAA TTG AGC GGC CTC GGC ACC GGG ATT CTG AT -3'

CS.s12

5'- TTT AAC CGC CAA GGG GAT TAC TCC CTA GTC -3'

CS.s23

5'- GCC GCT CAA TTG GTC GTA GAC AGC TCT AGC AC -3'

IS\textsuperscript{30}

5'- TTT AAC CGC CAA GGG GAT TAC TCC CTA GTC -3'

IS\textsuperscript{98}

5'- ATC AGA ATC CCG GTG CCG AGG CCG CTC AAT TGG TCG TAG ACA
GCT CTA GCA TTG CTT AAA TGT ACG TAC GCG CTG TCA TAT GAG TTT
TAA CCG CCA AG -3'
Figure 3.15 Ligation scheme for LS\textsubscript{U\textsubscript{110}} and CS. Nomenclature of each component and scaffold strand is indicated.
3.7 References


60. Tims HS, Gurunathan K, Levitus M, Widom J (2011) Dynamics of nucleosome


release of H3 and H4 from the nucleosome. *Biochemistry* 43(9):2359–2372.


Chapter 4

Global Repair Profile of Human Alkyladenine DNA Glycosylase on Nucleosomes Reveals DNA Packaging Effects

4.1 Abstract

Alkyladenine DNA glycosylase (AAG) is the only known human glycosylase capable of excising alkylated purines from DNA, including the highly mutagenic 1,N\(^6\)-ethenoadenine (εA) lesion. Here, we examine the ability of AAG to excise εA from a nucleosome core particle (NCP), which is the primary repeating unit of DNA packaging in eukaryotes. Using chemical synthesis techniques, we assembled a global population of NCPs in which A is replaced with εA. While each NCP contains no more than one εA lesion, the total population contains εA in 49 distinct geometric positions. Using this global εA-containing NCP system, we obtained kinetic parameters of AAG throughout the NCP architecture. We observed monophasic reaction kinetics across the NCP, but varying amounts of AAG excision. AAG activity is correlated with solution accessibility and local histone architecture. Notably, we identified some highly solution-accessible lesions that are not repaired well, and an increase in repair within the region of asymmetric unwrapping of the nucleosomal DNA end. These observations support in vivo work and provide molecular-level insight into the relationship between repair and NCP architecture.

4.2 Introduction

Despite being the code of life, DNA has a physiochemical composition that makes it susceptible to modification and decomposition by endogenous and exogenous sources (1). Failure to rectify these modifications can result in mutagenic consequences, such as cancer (2). The repair of nucleobase lesions can be accomplished with the base excision repair (BER) pathway. BER is initiated by a glycosylase enzyme specialized
Figure 4.1 Representations of εA:T bp, NCP, and DNA sequence. (A) An A:T bp and mutagenic εA:T bp. (B) Representation of an NCP. (C) The 145 bp Widom 601 duplex. Locations of A in sequence are highlighted in the I strand (green) and J strand (blue). Base pairs are numbered starting from the 5’-end of the I strand, with the J strand nucleobase indicated by a negative.

for removing a particular lesion (3). Most glycosylases use a base-flipping mechanism to extrude the lesion from the helix and catalyze $N$-glycosidic bond cleavage, resulting in an abasic site (4). Interestingly, alkyladenine DNA glycosylase (AAG) is the only known human glycosylase that removes alkylated nucleobases, including 3-methyladenine, 7-methyl-guanine, and 1,6-etheno-adenine (εA) (5). Specifically, εA can be generated by environmental exposure to vinyl chloride, or by attack of endogenous lipid peroxidation products (Figure 4.1A) (6). The current study focuses on the removal of εA by AAG.

Many prior studies characterized glycosylase activity using short oligonucleotide duplex substrates. However, eukaryotic DNA is highly packaged, presenting a more complicated environment for repair in vivo. The primary repeating unit of chromatin is the nucleosome core particle (NCP) (Figure 4.1B) with 75%-90% of genomic DNA packaged as such (7). NCPs consist of 145-147 bp of DNA wrapped in
~1.7 left-handed superhelical turns around an octameric core of histone proteins (8). The histone core contains two copies of each histone (H2A, H2B, H3, and H4), with 2-fold rotational symmetry about a dyad axis. As a consequence of wrapping around the histone core, the DNA becomes structurally distorted and locally stretched (9). The presence of the histones also means that there are regions of DNA that are sterically blocked while other regions are physically more accessible to DNA-binding proteins (10).

The geometric position of a nucleobase in an NCP can be described in two ways: the rotational position and the translational position. Rotational positioning of a nucleobase refers to its helical orientation; i.e., if the lesion is facing outward toward solution, or inward toward the histone core (Figure 1.4D in §1.5). Translational positioning refers to the location of the nucleobase relative to the dyad axis (Figure 1.4B in §1.5). It has been shown that nucleobases located toward the ends of the DNA have increased solution accessibility, compared with those near the dyad region, because of transient and spontaneous DNA unwrapping (10, 11).

Previous studies have shown that geometric positioning of a lesion and local histone environment can modulate glycosylase activity on an NCP (12). Most of these previous studies created NCPs containing a site-specific lesion. To obtain information about activity in other NCP regions, additional site-specific lesion-containing NCPs were created with the lesion in a different geometric position. These initial experiments showed that each local histone environment had varying effects on glycosylase activity beyond solution accessibility. Because every nucleobase in an NCP inherently has a unique microenvironment, no lesion location can be used as a representative of repair in a packaged DNA context. Furthermore, although studies using site-specific lesions in NCPs are informative, it would be time-consuming and impractical to study DNA repair across the entirety of the nuanced NCP architecture containing 145-147
bp in this manner.

With these considerations in mind, we recently designed a global system for simultaneous study of multiple lesion positions, and therefore probing of multiple NCP microenvironments, to describe glycosylase activity across an NCP (13). With the global system, this chapter examines the ability of AAG to excise εA from NCP at a variety of geometric positions, and discusses the structural nuances of NCP that are revealed by AAG repair profile.

4.3 Materials and Methods

4.3.1 Global εA-containing 145 mer oligonucleotide synthesis and purification.

All oligonucleotides used in this study were synthesized on a MerMade 4 DNA synthesizer (BioAutomation), and the sequences for them are shown in §4.6. All phosphoramidites and reagents were purchased from Glen Research. We utilized the 145 bp Widom 601 nucleosome positioning sequence for duplex and NCP assembly (14). Base pairs are numbered starting from the 5’-end of the I strand, with the J strand designated with a negative (-) (Figure 4.1C). The 145 mer oligonucleotides containing global εA were synthesized on 1400Å controlled pore glass beads using phosphoramidites with ultramild protecting groups and deprotected according to the manufacturer’s specifications. εA was substituted for A throughout the DNA on either the I strand or J strand of the Widom 601 sequence to ultimately have at most one εA:T bp per DUP. To accomplish this, we used methods similar to our recent report (13). Briefly, a mixture of εA and A phosphoramidites was used during the synthesis with the molar ratio determined by the Poisson distribution (λ = 0.355), such that 95%
of the DNA population contains either 0 or 1 εA lesion per 145 mer oligonucleotide. The final trityl group was removed on the synthesizer. The DNA was cleaved from the beads by incubation in NH₄OH at room temperature for 2 h and purified by 8% denaturing PAGE run for 5 h at 80 W. A small portion of the synthesized DNA was radiolabeled and was loaded in lanes adjacent to where the unlabeled DNA was loaded. The region of the gel that co-migrated with a 145 mer standard was excised and “crushed and soaked” (15) overnight in 15 mL elution buffer (300 mM sodium acetate [pH 8.0], 1 mM EDTA) with gentle shaking (80 rpm) at 37 °C. The buffer containing the eluted DNA was passed through a 0.22 µm cellulose acetate syringe filter and was concentrated and desalted with many ethanol precipitations. The concentrations of the PAGE-purified oligonucleotides were determined by their absorbance at 260 nm using molar extinction coefficients for the undamaged Widom 601 sequences obtained using the OligoAnalyzer tool on www.idtdna.com.

The incorporation of εA at A sites was confirmed by cleavage by AAG. Specifically, the εA-containing 145 mer strand was 5’-radiolabeled using T4 kinase (New England Biolabs) and γ-32P-ATP (Perkin Elmer) and annealed to its respective undamaged complement to form duplex. The εA-containing duplex (DUP) was then incubated with AAG for 1 h at 37 °C, followed by the addition of equal volume of 1 M NaOH and incubation at 90 °C for 3 min. Samples were supplemented with 40 µL co-precipitation agent (0.5 mg/mL tRNA in 300 mM NaOAc [pH 8.0], 1 mM EDTA) and subsequently desalted with two ethanol precipitations. 12% denaturing PAGE gel was then utilized to visualize cleavage at εA created by AAG treatment.
4.3.2 Ligation strategy to synthesize undamaged 145 mer oligonucleotides complementary to εA-containing strands.

The 145 mer oligonucleotides complementary to the εA-containing oligonucleotides were prepared by ligating three short component oligonucleotides (see Figure 4.15 in §4.6 for detailed ligation scheme). The component oligonucleotides for ligation were synthesized using standard phosphoramidite protecting groups with the final trityl group retained for reverse-phase HPLC purification at 90 °C (Agilent PLRP-S column, 250 mm × 4.6 mm; A = 100 mM triethylammonium acetate [TEAA] in 5% aqueous MeCN, B = 100 mM TEAA in MeCN; 5:95 to 35:65 A:B over 30 min, 35:65 to 5:95 A:B over 5 min at 1 mL/min). The trityl group was then removed with incubation in 20% v/v aqueous glacial acetic acid for 1 h at room temperature, followed by a second HPLC purification at 90 °C (Agilent PLRP-S column, 250 mm × 4.6 mm; A = 100 mM triethylammonium acetate [TEAA] in 5% aqueous MeCN, B = 100 mM TEAA in MeCN; 0:100 to 15:85 A:B over 35 min, 15:85 to 35:65 A:B over 5 min at 1 mL/min). Electrospray ionization mass spectrometry was used to verify the identity of the component oligonucleotides.

Component oligonucleotides were used in ligation. 5 nmol of each component oligonucleotide 2 and 3 were 5'-phosphorylated in the presence of 2 mM ATP using T4 kinase (New England Biolabs). Phosphorylated component oligonucleotides were combined in equal amounts with component 1 and a 5% excess of two scaffolding oligonucleotides, s12 and s23, in annealing buffer (10 mM Tris-HCl pH 8.0, 50 mM NaCl, 1 mM EDTA). The strands were annealed by heating to 90 °C for 5 min followed by cooling to room temperature at a rate of 1 °C/min. The annealed component oligonucleotides were ligated together overnight at room temperature using T4 ligase (New England Biolabs). The ligated 145 mer oligonucleotides were purified
by denaturing PAGE as described above for the εA-containing 145 mer.

4.3.3 Formation of global εA-containing duplex.

The 145 mer oligonucleotides containing εA lesions were 5’- radiolabeled for visualization. The global εA-containing I strand was annealed with the J strand 145 mer that lacked εA lesions. Similarly, the J strand oligonucleotide containing εA lesions was annealed with the undamaged I strand. DUP was formed by mixing the radiolabeled εA-containing strand with the complementary strand in a 1:1.07 ratio in annealing buffer and heating to 90 °C for 5 min followed by cooling to room temperature at a rate of 1 °C/min. DUP formation was confirmed on a 7% native PAGE (60:1 acrylamide : bisacrylamide; 0.25X TBE). The samples were loaded using 5% (v/v) glycerol in 10 mM Tris-HCl (pH 7.5), 1 mM EDTA and run for 3 h at 160 V at 4 °C. Only DUP containing 5% single-strand DNA were used in further studies.

4.3.4 Reconstitution of global εA nucleosome core particles (NCPs).

Recombinant Xenopus laevis histones were individually expressed and purified, and subsequently assembled into octamers (16, 17). NCPs were reconstituted by dialyzing the radiolabeled εA-containing DUP population and histone octamer together via salt gradient, as described previously (13). Briefly, a 10% molar excess of histone octamer was gently added to 25 μL of 1 μM radiolabeled εA-containing 145 bp DUP in buffer (10 mM Tris-HCl [pH 7.5], 1 mM EDTA, 1 mM dithiothreitol (DTT), 2 M NaCl, 500 μg/mL BSA) in a Slide-a-Lyzer dialysis device (0.1 mL capacity, 3.5 kDa MWCO; Thermo Fisher Scientific). The dialysis device started in a buffer of 10 mM Tris-HCl (pH 7.5), 1 mM EDTA, 1 mM DTT, 2 M NaCl at 4 °C. At 1 h intervals, the device
was placed in analogous buffers containing decreasing concentrations of NaCl (1.2 M, 1.0 M, 0.6 M, 0 M). The final dialysis in 0 M NaCl proceeded for 3 h before the reconstitution was filtered with a 0.22 μm cellulose acetate centrifuge tube filter (Corning Costar) to remove precipitates. NCP formation and relative purity were analyzed using a 7% native PAGE (60:1 acrylamide: bisacrylamide; 0.25x TBE) run for 3 h at 160 V in 4 °C. Only NCPs containing ≤ 5% DUP were used in further studies. In this work, the upper limit of contaminating DUP used in NCP experiments is the lowest amount of εA excision observed at a single site. Site −56 has an average product accumulation of 4%. Therefore, DUP cannot account for more than 4% of the reported NCP product.

We tested for bias of εA incorporation into NCPs to address the concern of εA enrichment or depletion as a function of εA location. We reconstituted global εA-containing DUP into NCPs and then subsequently removed the histone proteins by extraction with an equal volume of 25:24:1 phenol:chloroform:isoamyl alcohol (PCI) and subsequent ethanol precipitation to create “freed DUP”. The freed DUP was incubated with a 30-fold excess of AAG for 2 h at 37 °C in 20 mM Tris-HCl (pH 7.6), 50 mM NaCl, 150 mM KCl, 1 mM EDTA, 1 mM DTT, 200 μg/mL BSA. Samples were quenched with the addition of an equal volume of 1 M NaOH supplemented with internal references. An equal volume of PCI was added and the aqueous layer was subjected to two ethanol precipitations. The samples were prepared in a 1:1 mixture of formamide and water for separation on a 12% denaturing PAGE to resolve the 5′-end of εA positions (ranging from 28-68). Quantitation of band intensity was determined by SAFA (18). The density measurements for all lanes was normalized to the internal references; the 23 mer internal reference was used for the 5′-ends of the sequences (ranging from fragment lengths 9-68) and the 92 mer internal reference was used for the 3′-ends of the sequences (ranging from fragment lengths 70 to 132).
The amount of background damage from the -E sample was subtracted from AAG-treated samples. The resulting amount of cleavage due to AAG activity observed in the freed DUP substrate was compared with analogous conditions of the same DUP population that had never been incorporated into NCPs. The ratio of εA excision in the freed DUP substrate was compared with εA excision in the unincorporated DUP to quantify equal distribution of εA lesions across the NCP populations. A ratio of 1 indicates that the amount of cleavage in the freed DUP is equal to that in the unincorporated DUP, and therefore present in equal amounts in both DUP substrates.

4.3.5 **Hydroxyl radical footprinting.**

HRF was carried out using previously-published conditions to establish relative solution accessibility of nucleobase positions in NCPs under single-hit conditions (9, 19). Briefly, 7.5 µL of each 1 mM Fe(II)-EDTA, 10 mM sodium ascorbate, and 0.12% w/v aqueous hydrogen peroxide were gently combined with 5 pmol NCPs in a total of 52.5 µL buffer (10 mM Tris-HCl [pH 7.5], 1 mM EDTA). The reaction proceeded at room temperature for 10 min in the dark and was quenched with the addition of 16 µL 50 mM EDTA in 25% v/v glycerol. The quenched sample was immediately loaded onto a 7% native PAGE (60:1 acrylamide: bisacrylamide; 0.25X TBE) and run for 3 h at 160 V at 4 °C. Gel bands containing the NCP species were excised and NCPs were eluted into buffer (0.3 M NaOAc, 1 mM Tris-HCl [pH 8.0], 1 mM EDTA) for 16-20 h at 37 °C with gentle shaking (80 rpm). The NCP eluent was concentrated using a centrifugal concentrator (Sartorius Vivaspin Turbo 15, 5 kDa MWCO) and filtered using a 0.22 µm cellulose acetate centrifuge tube filter (Corning Costar). The sample was extracted twice with equal volume additions of PCI to remove histone
proteins, and the resulting aqueous phase was concentrated by SpeedVac evaporation. Following the addition of 40 μL co-precipitation agent (0.5 mg/mL tRNA in 300 mM NaOAc [pH 8.0], 1 mM EDTA), samples were desalted with two ethanol precipitations. Samples were prepared in a 1:1 mixture of formamide and water for denaturing PAGE. Cleavage fragments were resolved after splitting the sample into two halves. Each half was loaded onto separate denaturing PAGE (12% gel for 5′-end fragments, 8% for 3′-end fragments) (20, 21). Denaturing PAGE were visualized by phosphorimagery (Bio-Rad PharosFX). Bands resulting from HRF cleavage were quantitated using SAFA gel analysis software (18). The more solution-accessible a nucleobase position is to hydrogen abstraction, the stronger the intensity of the corresponding cleavage fragment. Categorization of solution accessibilities of nucleobases were determined as a ratio of band intensity at a given position relative to the highest band intensity within the helical turn. Highly solution-accessible (HIGH) positions were defined as those with a ratio greater than 0.7; medium solution-accessible (MID) positions with a ratio range from 0.3-0.7; solution-inaccessible (LOW) positions were defined as those with a ratio less than 0.3.

4.3.6 DNase I footprinting of NCPs.

NCPs were reconstituted using DUP containing no εA lesions (in 10 mM Tris-HCl [pH 7.5], 1 mM EDTA, 1 mM DTT, 2 M NaCl, 500 μg/mL BSA) and were treated with varying amounts of DNase I (New England Biolabs). 2.5 pmol NCP (containing < 2% DUP determined by native PAGE) were incubated with 0.02 and 0.002 U DNase I at 37 °C for 5 min in 1x DNase I buffer (10 mM Tris-HCl [pH 7.6], 7.5 mM MgCl2, 0.5 mM CaCl2) supplemented with an additional 10 mM MgCl2 for a final
concentration of 17.5 mM MgCl₂. The reaction was quenched with the addition of one-third the reaction volume of 50 mM EDTA–0.5% sodium dodecyl sulfate (SDS)–0.2 mg/mL proteinase K, followed by incubation at 50°C for 2 h. Protein extraction was performed with equal volume additions of PCI, followed by two ethanol precipitations.

As a control, DNase I footprinting was performed on DUP in analogous buffer conditions, but incubated at 37 °C for 2 min. The reaction was quenched with the addition of half the reaction volume of 100 mM EDTA-1 mg/mL calf-thymus DNA. Samples were ethanol precipitated twice.

All samples were prepared in a 1:1 mixture of formamide and water for denaturing PAGE similar to those described for HRF above. Denaturing PAGE were visualized by phosphorimagery (Bio-Rad PharosFX). Bands resulting from DNase I cleavage were quantitated using SAFA gel analysis software. Similar to HRF, the more solution accessible a nucleobase position is for cleavage, the stronger the intensity of the corresponding cleavage fragment.

4.3.7 AAG kinetics on global εA-containing DNA.

AAG was purchased from New England Biolabs. The total concentration of AAG was determined by the Bradford method using bovine γ-globulin standards (Bio-Rad Laboratories). All AAG concentrations given below or in figure captions are total enzyme concentrations. Reactions were initiated after the DNA substrate and AAG were pre-equilibrated at 37 °C for 2 min by combining an equal volume of 26.6 nM DNA substrate (DUP or NCP) and 0.8 μM AAG for a final experimental sample of 13.3 nM DNA and 0.4 μM AAG in 20 mM Tris-HCl (pH 7.6), 50 mM NaCl, 150 mM KCl, 1 mM EDTA, 1 mM DTT, 200 μg/mL BSA. The reactions were incubated at 37 °C for 5, 15, 30, 60, 90, 120, or 180 min before they were quenched with an equal
volume of 1 M NaOH supplemented with the radiolabeled internal references and incubated at 90 °C for 3 min. Timepoints were initiated in such a way that all samples were quenched at the same time before immediate sample workup; this process was used to avoid extended incubation of εA lesions in NaOH after quenching which could result in degradation and false-positive results of AAG-catalyzed cleavage (22, 23). After incubation with NaOH, an equal volume of PCI was added followed by 40 µL of co-precipitation agent (0.5 mg/mL tRNA in 300 mM NaOAc [pH 8.0], 1 mM EDTA) and the samples were subsequently desalted with two ethanol precipitations. The -E sample was not incubated with AAG but instead an equivalent volume of reaction buffer was added and incubated at 37 °C for 180 min, quenched, and subsequently worked up like the rest of the samples. Any preexisting damage, such as depurination, or incidental damage due to experimental conditions or workup was revealed by this NaOH treated sample, which was used for a background subtraction for AAG-treated samples. Therefore, the amount of εA excision reported is strictly due to AAG’s enzymatic activity.

After the final ethanol precipitation all samples were resuspended in a 1:1 mixture of formamide and water and split in half. Each half was loaded onto separate denaturing PAGE (12% gel for 5’-end fragments; 8% for 3’-end fragments). The PAGE gels were visualized by phosphorimagery on a Bio-Rad PharosFX and quantified by SAFA. The internal references served as a loading control and also allowed us to account for loss of εA cleavage products due to sample workup, especially smaller fragments that have a decreased ability to precipitate (24). Accordingly, the band intensities in each lane were normalized to an internal reference. The 23 mer internal reference was used for the 5’-end fragments (ranging from 9 to 68) and the 92 mer internal reference was used for the 3’-ends fragments (ranging from 70 to 132). After
this normalization the band intensity in the -E lane was subtracted from the AAG-treated samples.

As a result of the global lesion substitution, <1% of the DNA population is available as a substrate at each εA position. Therefore, the exposure time needed for phosphorimagery leads to overexposure of the parent band which then cannot be accurately quantified and used as a reference point to determine the amount of cleavage product at each εA. To address this technical challenge, we defined the amount of product observed with the DUP control after 180 min as the theoretical maximum amount of product at a given lesion position (DUP$_{180}$). Therefore, product accumulation at time $t$ is represented as a ratio relative to DUP$_{180}$. In DUP, product accumulation at time $t$ is represented as DUP$_t$/DUP$_{180}$, where DUP fraction product is normalized to 1.0 for the 180 min sample. In NCPs, product accumulation at time $t$ is represented as NCP$_t$/DUP$_{180}$, where a ratio of 1.0 indicates the εA is removed completely, as in DUP. Replicate data at each observed εA position were averaged for every timepoint, and then fit to a single exponential model:

$$y(t) = y_{max}(1 - e^{-k_{obs}t}),$$

(4.1)

where $y_{max}$ is the amplitude of the reaction’s fit, $k_{obs}$ is the observed rate constant, and $t$ is time. Error bars shown are calculated standard errors (SE) from 3 to 10 replicates for each εA position at each timepoint:

$$SE = \frac{\sigma}{\sqrt{n}},$$

(4.2)

where $\sigma$ is the standard deviation of the population and $n$ is the number of replicates.

We confirmed that AAG retains full activity during the 180 min time course.
The activity of AAG pre-incubated for 180 min (3 h sample) at 37 °C under experimental conditions was compared to that of AAG directly added (0 h sample) to global εA-containing DUP. After addition, the reactions proceeded for 1 h at 37 °C before quenching with NaOH supplemented with radiolabeled internal references and sample workup as described above. Cleavage products resulting from εA excision were resolved on a 12% denaturing PAGE to resolve the 5′-end fragments and quantified by SAFA (18). After normalization using the 53 mer internal reference and background subtraction of the -E data the ratio of product formation for the 180 min sample relative to the 0 h sample was determined. A ratio of 1 indicates that after 180 min at 37 °C AAG retains full activity.

4.3.8 Molecular modeling of NCP

A crystal structure modeling the Widom 601 sequence wrapped around a *Xenopus laevis* histone octamer with histone tails was created using the “merge” function in PyMOL. The NCP crystal structure containing the Widom 601 sequence but which lacks histone tails (PDB ID: 3lz0) was merged with an NCP crystal structure containing the α-satellite DNA sequence but containing the histone tails (PDB ID: 1kx5) by aligning the overlapping histone protein amino acid sequences. The Widom 601 duplex and the histone octamer containing the histone tails were combined and merged to create a new, separate PyMOL object.
4.4 Results and Discussion

4.4.1 AAG activity on global εA-containing duplex DNA

Global lesion-containing DNA is prepared via chemical synthesis, using a building block mixture of the lesion of interest and the unmodified nucleobase that maximizes the number of strands containing 0 or 1 lesion throughout the sequence. Reconstitution of NCPs using global lesion-containing duplex and histone octamer yields a population of NCPs with the lesion positioned in a variety of well-defined translational and rotational positions. Here, we utilize the global system to substitute εA for A in either the I or J strand of the Widom 601 NCP (25) to study excision by AAG in a total of 49 distinct geometric positions (see Figure 4.1C and Figure 4.2). Whereas our previous work qualitatively described the repair of 8-oxo-7,8-dihydroguanine lesions in NCPs, our current work elevates the global system to provide kinetic parameters to describe εA excision by AAG in duplex controls and NCPs.

We first measured εA excision by AAG, as a function of time, in a 145 bp global duplex control (DUP) (see Figure 4.3A and 4.4A). DUP was assembled with one εA-containing 145 mer annealed to the undamaged complementary strand. Because of the small amount of each lesion-containing substrate, relative to the total DNA population, we defined the amount of excised product observed after 180 min as the maximal product accumulation at a given lesion site. We chose an 180 min time course, based on previous experiments and the slow activity of AAG (5, 26), and confirmed that AAG retains full activity over the 180 min (Figure 4.5). Therefore, product accumulation at time $t$ is represented as a ratio relative to product accumulation at 180 min (DUP$_t$/DUP$_{180}$) (Figure 4.6). Monophasic fits of reaction progress at each lesion site in DUP (Figure 4.3B, closed symbols) determined the observed rate ($k_{obs}$) of AAG excision at the 49 εA positions in DUP, which ranges from 0.030
Figure 4.2 Confirmation of εA substitutions for A in Widom 601 I and J strands. The εA was incorporated separately in the (A) I and (B) J strands. The L lanes display a sequence ladder created using Maxam-Gilbert reactions using Widom 601 sequences lacking εA lesions. NT lanes show untreated I or J strand. The +E lanes are DNA duplexes containing εA in either the I or J strand treated with AAG for 1 h at 37 °C and quenched with the addition of NaOH to create strand breaks at AAG-generated abasic sites. In both PAGE images, the parent band was partially cropped for clarity.
to 0.091 min\(^{-1}\) (Table 4.1). These \(k_{\text{obs}}\) values are comparable to previous reports for site-specific εA lesions in both short oligonucleotide duplexes (5, 26) and 145 bp duplex (17).

### 4.4.2 AAG activity on global εA-containing NCPs

We next prepared global εA-containing NCP populations to examine AAG activity in packaged DNA. Successful formation of NCPs is indicated by native PAGE gel (Figure 4.7). We previously reported (13, 17, 27) extensive control experiments to determine strategies that minimize, to the greatest extent possible, unincorporated DNA in NCP preparations. Only NCP preparations with \(\leq 5\%\) DUP are used in experiments. Product accumulation in the NCP was determined at time \(t\) as a ratio relative to maximal product accumulation, \(\text{NCP}_t / \text{DUP}_{180}\) (see Figure 4.3A and C for data of the J strand, and Figure 4.4 for data of the I strand). Unlike AAG excision from DUP, not all εA lesions in NCPs were fully excised by 180 min; 30/49 sites had <30\% excision (Figure 4.8, green and blue bars). Suppressed activity, relative to DUP, indicates that the presence of the histone octamer interferes with AAG activity. We verified that all DNA containing εA lesions formed NCPs and there was no bias in NCP formation based on the geometric position of the lesion (Figure 4.9); therefore, a lack of excision from NCPs is due to a lack of AAG activity.
Figure 4.3 Single-turnover kinetic time courses of εA excision by AAG on the J strand. (A) Representative PAGE gels showing εA excision on the J strand of DUP (left) and NCP (right). DNA samples only treated with NaOH (-E lanes) show incidental damage from experimental conditions and sample workup. Two internal references, a 23 mer (*) and 92 mer (**), used for quantitation are indicated. Four separate gels are shown, indicated by the spaces between the images, with each gel cropped for alignment. The top two panels are aligned to show DUP and NCP data for positions –8 to –76. The bottom two panels are aligned to show data for positions –85 to –135. (B) Selected kinetic fits showing examples of low (site –87), medium (site –128), and high (position –133) amounts of εA excision on DUP (closed symbols) and NCPs (open symbols). (C) Quantitation of εA excision over time on the J strand in NCPs. Error bars are calculated standard errors (n=3-10). Analogous data for the I strand in NCPs is shown in Figure 4.4.
Figure 4.4 Single-turnover kinetic time courses of εA excision by AAG on the I strand. (A) Representative PAGE gels showing εA excision on the J strand of DUP (left) and NCP (right). DNA samples only treated with NaOH (-E lanes) show incidental damage from experimental conditions and sample workup. Two internal references, a 23 mer (*) and 92 mer (**), used for quantitation are indicated. Four separate gels are shown, indicated by the spaces between the images, with each gel cropped for alignment. The top two panels are aligned to show DUP and NCP data for positions -8 to -76. The bottom two panels are aligned to show data for positions -85 to -135. (B) Selected kinetic fits showing examples of low (site -87), medium (site -128), and high (position -133) amounts of εA excision on DUP (closed symbols) and NCPs (open symbols). (C) Quantitation of εA excision over time on the J strand in NCPs. Error bars are calculated standard errors (n=3-10).
Figure 4.5 AAG retains full activity for 180 min. (A) AAG was either added directly (0 h lane) or pre-incubated at 37 °C under experimental conditions for 180 min (3 h lane) before incubation with global εAI DUP. After AAG addition, the reaction proceeded for 1 h at 37 °C before quenching with NaOH and sample workup for loading on denaturing PAGE. Three internal references, a 23 mer (*), 92 mer (**), and 53 mer (***) used for quantitation are indicated. The N lane shows the DUP sample without any treatment. The -E lane is quenched with NaOH to show any damage due to experimental conditions and workup. (B) Quantification of AAG activity at various positions in DUP, shown as a ratio of εA excision seen in the 3 h pre-incubation sample to the amount of εA excision seen in the 0 h sample. A ratio of 1 indicates that pre-incubation of AAG for 3 h before reaction does not influence its activity.
Figure 4.6 Quantitation of εA excision over time in global DUP. Lesions removed from the (A) I strand and (B) J strand at the position locations indicated. Product accumulation at time \( t \) is represented as a ratio relative to product accumulation at 180 min (DUP\(_ t \)/DUP\(_ {180} \)), the defined maximal product. Error bars are calculated standard errors (n=3-10).
Table 4.1. \( k_{\text{obs}} \) values for \( \epsilon A \) excision by AAG.\textsuperscript{a}

<table>
<thead>
<tr>
<th>Base position</th>
<th>( k_{\text{obs}} ) (min(^{-1}))</th>
<th>Frac. pdt</th>
<th>( R^2 ) fit</th>
<th>( k_{\text{obs}} ) (min(^{-1}))</th>
<th>Frac. pdt</th>
<th>( R^2 ) fit</th>
</tr>
</thead>
<tbody>
<tr>
<td>19</td>
<td>0.057</td>
<td>0.965</td>
<td>0.991</td>
<td>0.044</td>
<td>0.133</td>
<td>0.817</td>
</tr>
<tr>
<td>-26</td>
<td>0.060</td>
<td>0.965</td>
<td>0.955</td>
<td>n.d.\textsuperscript{b}</td>
<td>&lt; 0.18</td>
<td>n.d.</td>
</tr>
<tr>
<td>28</td>
<td>0.037</td>
<td>0.990</td>
<td>0.977</td>
<td>0.019</td>
<td>0.157</td>
<td>0.95</td>
</tr>
<tr>
<td>29</td>
<td>0.067</td>
<td>0.970</td>
<td>0.977</td>
<td>0.054</td>
<td>0.139</td>
<td>0.932</td>
</tr>
<tr>
<td>-30</td>
<td>0.086</td>
<td>0.954</td>
<td>0.933</td>
<td>0.017</td>
<td>0.435</td>
<td>0.705</td>
</tr>
<tr>
<td>-31</td>
<td>0.038</td>
<td>0.979</td>
<td>0.967</td>
<td>0.009</td>
<td>1.092</td>
<td>0.949</td>
</tr>
<tr>
<td>-34</td>
<td>0.089</td>
<td>0.930</td>
<td>0.861</td>
<td>n.d.</td>
<td>&lt; 0.240</td>
<td>n.d.</td>
</tr>
<tr>
<td>-37</td>
<td>0.054</td>
<td>0.988</td>
<td>0.992</td>
<td>n.d.</td>
<td>&lt; 0.240</td>
<td>n.d.</td>
</tr>
<tr>
<td>38</td>
<td>0.054</td>
<td>0.982</td>
<td>0.979</td>
<td>0.073</td>
<td>0.146</td>
<td>0.884</td>
</tr>
<tr>
<td>40</td>
<td>0.070</td>
<td>0.957</td>
<td>0.970</td>
<td>0.087</td>
<td>0.167</td>
<td>0.727</td>
</tr>
<tr>
<td>42</td>
<td>0.036</td>
<td>0.930</td>
<td>0.964</td>
<td>0.027</td>
<td>0.880</td>
<td>0.966</td>
</tr>
<tr>
<td>-45</td>
<td>0.038</td>
<td>0.991</td>
<td>0.983</td>
<td>n.d.</td>
<td>&lt; 0.120</td>
<td>n.d.</td>
</tr>
<tr>
<td>-47</td>
<td>0.040</td>
<td>0.984</td>
<td>0.988</td>
<td>n.d.</td>
<td>&lt; 0.120</td>
<td>n.d.</td>
</tr>
<tr>
<td>48</td>
<td>0.051</td>
<td>0.958</td>
<td>0.963</td>
<td>0.070</td>
<td>0.171</td>
<td>0.842</td>
</tr>
<tr>
<td>51</td>
<td>0.055</td>
<td>0.943</td>
<td>0.970</td>
<td>0.052</td>
<td>0.443</td>
<td>0.785</td>
</tr>
<tr>
<td>-56</td>
<td>0.046</td>
<td>0.990</td>
<td>0.995</td>
<td>n.d.</td>
<td>&lt; 0.055</td>
<td>n.d.</td>
</tr>
<tr>
<td>-57</td>
<td>0.034</td>
<td>1.020</td>
<td>0.992</td>
<td>n.d.</td>
<td>&lt; 0.065</td>
<td>n.d.</td>
</tr>
<tr>
<td>58</td>
<td>0.050</td>
<td>0.965</td>
<td>0.974</td>
<td>n.d.</td>
<td>&lt; 0.220</td>
<td>n.d.</td>
</tr>
<tr>
<td>59</td>
<td>0.052</td>
<td>0.955</td>
<td>0.977</td>
<td>0.073</td>
<td>0.156</td>
<td>0.907</td>
</tr>
</tbody>
</table>
Table 4.1 (cont’d)

<table>
<thead>
<tr>
<th>Base position</th>
<th>DUP</th>
<th>NCP</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$k_{\text{obs}}$ (min$^{-1}$)</td>
<td>Frac. pdt</td>
</tr>
<tr>
<td>60</td>
<td>0.059</td>
<td>0.960</td>
</tr>
<tr>
<td>64</td>
<td>0.046</td>
<td>0.964</td>
</tr>
<tr>
<td>-67</td>
<td>0.039</td>
<td>0.952</td>
</tr>
<tr>
<td>68</td>
<td>0.053</td>
<td>0.990</td>
</tr>
<tr>
<td>-74</td>
<td>0.031</td>
<td>0.994</td>
</tr>
<tr>
<td>-76</td>
<td>0.068</td>
<td>0.948</td>
</tr>
<tr>
<td>-85</td>
<td>0.088</td>
<td>0.951</td>
</tr>
<tr>
<td>-86</td>
<td>0.065</td>
<td>1.007</td>
</tr>
<tr>
<td>-87</td>
<td>0.070</td>
<td>1.007</td>
</tr>
<tr>
<td>-88</td>
<td>0.056</td>
<td>1.007</td>
</tr>
<tr>
<td>89</td>
<td>0.048</td>
<td>1.005</td>
</tr>
<tr>
<td>90</td>
<td>0.060</td>
<td>1.001</td>
</tr>
<tr>
<td>96</td>
<td>0.048</td>
<td>1.000</td>
</tr>
<tr>
<td>97</td>
<td>0.059</td>
<td>0.983</td>
</tr>
<tr>
<td>102</td>
<td>0.076</td>
<td>0.983</td>
</tr>
<tr>
<td>-103</td>
<td>0.091</td>
<td>0.980</td>
</tr>
<tr>
<td>-104</td>
<td>0.043</td>
<td>0.993</td>
</tr>
<tr>
<td>105</td>
<td>0.041</td>
<td>1.013</td>
</tr>
<tr>
<td>-107</td>
<td>0.065</td>
<td>0.977</td>
</tr>
</tbody>
</table>
Table 4.1 (cont’d)

<table>
<thead>
<tr>
<th>Base position</th>
<th>DUP</th>
<th>NCP</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$k_{\text{obs}}$ (min$^{-1}$)</td>
<td>Frac. pdt</td>
</tr>
<tr>
<td>–111</td>
<td>0.054</td>
<td>0.969</td>
</tr>
<tr>
<td>112</td>
<td>0.046</td>
<td>0.963</td>
</tr>
<tr>
<td>–114</td>
<td>0.079</td>
<td>0.957</td>
</tr>
<tr>
<td>–116</td>
<td>0.071</td>
<td>0.964</td>
</tr>
<tr>
<td>119</td>
<td>0.059</td>
<td>0.958</td>
</tr>
<tr>
<td>123</td>
<td>0.068</td>
<td>0.945</td>
</tr>
<tr>
<td>–126</td>
<td>0.049</td>
<td>0.964</td>
</tr>
<tr>
<td>–128</td>
<td>0.080</td>
<td>0.976</td>
</tr>
<tr>
<td>–133</td>
<td>0.031</td>
<td>0.965</td>
</tr>
<tr>
<td>–135</td>
<td>0.033</td>
<td>0.982</td>
</tr>
<tr>
<td>–137</td>
<td>0.043</td>
<td>0.962</td>
</tr>
</tbody>
</table>

$^a$Product accumulation over time was characterized by monophasic kinetic fits by Kaledagrap to yield $k_{\text{obs}}$ and % product values for AAG excising εA at different positions in I and J strands.

$^b$Fields displaying “n.d.” were unable to be determined based on the small amounts of product accumulation observed.

We next obtained $k_{\text{obs}}$ describing AAG excision of εA from the global NCPs. As observed for DUP, excision of εA across NCPs displayed monophasic behavior (Figure 4.3B, open symbol). In addition, $k_{\text{obs}}$ values across NCPs are comparable to those observed in DUP and to those obtained for a site-specific lesion in an NCP (Table 4.1) (17). However, different from results observed with DUP, significantly suppressed excision efficiency is observed at 30/49 sites. The monophasic behavior of AAG in DUP and NCP, paired with the varying amounts of εA excision in NCPs, suggests that there are subpopulations of NCPs that are conformationally competent for AAG activity, while others are not. The NCP population that is unrepai red by AAG is not due to the DNA substrates being chemically incompetent. Instead, the
Figure 4.7 Formation of global εA-containing DUP and NCPs. Representative native PAGE gels showing differential migrations of global εA-containing single-strand (ss), DUP, and NCPs on the (A) I strand and (B) J strand. When εA was in the I strand it was paired with the J strand lacking damage. When εA was in the J strand, it was paired with I strand lacking damage. Only DUP preparations having ≤ 5% single-strand DNA were used in further studies. Similarly, only NCP preparations having ≤ 5% DUP were used in further studies.

Figure 4.8 Relationship between AAG activity on and solution accessibility of εA positions in the NCP. The amount of AAG excision after 180 min is shown as bars at each εA position along the I strand (green) and J strand (blue). Error bars represent the standard error (n=3-10). The HRF profile characterizes relative reactivity as a function of base position and is depicted as gray area in the background.
Figure 4.9 Non-biased incorporation of εA-containing DNA into NCPs. NCPs were assembled with global εA-containing I strand DUP, and then separated from histone octamers using a PCI extraction to create “freed DUP”. AAG was incubated with freed DUP for 2 h at 37 °C and with DUP never incorporated into an NCP to compare differences in AAG activity. Different amounts of εA excision between the two samples would suggest biased incorporation of εA lesions during NCP reconstitution. (A) Denaturing PAGE gel of reactions for quantitative comparison. Two internal references, a 23 mer (*) and 53 mer (***) used for quantitation are indicated. –AAG lanes were only treated with NaOH to show any damage due to experimental conditions and workup. The freed DUP samples are indicated with +PCI designations. (B) Quantification of εA excision in each DUP substrate. A ratio of 1 indicates that the amount of cleavage in the freed DUP is equal to that in the unincorporated DUP.
Table 4.2. Solution accessibilities of adenine positions in NCPs.

<table>
<thead>
<tr>
<th>Solution accessibility</th>
<th>HRF ratio&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Positions</th>
</tr>
</thead>
<tbody>
<tr>
<td>LOW</td>
<td>&lt; 0.3</td>
<td>19, −26, 28, 29, 38, −45, 48, −56,</td>
</tr>
<tr>
<td></td>
<td></td>
<td>−57, 58, 59, 60, −67, 68, −76, −86,</td>
</tr>
<tr>
<td></td>
<td></td>
<td>−87, 88, 89, 90, −107, 112, 123, −128</td>
</tr>
<tr>
<td>MID</td>
<td>0.3 - 0.7</td>
<td>−34, −37, 40, −47, 51, −85, 102, −111,</td>
</tr>
<tr>
<td></td>
<td></td>
<td>−116, 119, −126, −137</td>
</tr>
<tr>
<td>HIGH</td>
<td>&gt; 0.7</td>
<td>−30, −31, 42, 64, −74, 96, 97, −103,</td>
</tr>
<tr>
<td></td>
<td></td>
<td>−104, 105, −114, −133, −135</td>
</tr>
</tbody>
</table>

<sup>a</sup>The HRF ratio is the ratio of band intensity at a given cleavage fragment in the HRF profile relative to the highest band intensity within the associated helical turn. This ratio represents the amount of solution accessibility of the nucleobase position relative to the most solution-accessible nucleobase position within the helical turn. LOW was defined as having a ratio of less than 0.3, MID as 0.3-0.7, and HIGH as more than 0.7.

population is not repaired, because of a structural impediment and/or conformation derived from the histone octamer, which renders the lesion inaccessible to AAG activity. In the cell, repair of these lesions might require chromatin remodelers and/or additional factors. Previous kinetic characterizations of human glycosylases OGG1 and UNG2 acting on site-specific lesion NCP systems revealed multiphasic kinetics (≥2 phases) for the removal of lesions in some positions, which was attributed to conformation changes in the NCP (27, 28). Considering the slower rate of glycosylase activity for AAG, compared to that of OGG1 and UNG2, the monophasic behavior observed across the NCP suggests that analogous conformational changes cannot be resolved on this time scale.
4.4.3 Relationships between AAG activity and solution accessibility of εA lesions in NCPs

Next, we evaluated the relationship between AAG activity and solution accessibility of each εA lesion. We performed hydroxyl radical footprinting (HRF) to characterize solution accessibility of nucleobases in global NCPs (Figure 4.10 and the shaded gray area in Figure 4.8). εA positions with high (HIGH), medium (MID), and low (LOW) solution accessibilities were grouped based on the ratios of accessibility to the maximal accessibility within a helical turn of the HRF profile (Table 4.2). Overall, we found that the extent of εA excision from NCPs by AAG is highly correlated with solution accessibility (Figure 4.8), similar to our previous observations of the global repair profiles of OGG1 (13) and UDG (29) on NCPs. Approximately half of the HIGH positions (i.e., sites 42, 64, −103, −104, 105, −114) had maximal εA excision (90%-100%). Conversely, minimal amounts of product were observed at most LOW positions, with no greater than 20% εA excision. MID positions had low to moderate εA excision, ranging from 10% to 60%. The broader range of εA excision observed at MID positions is likely due to varied NCP microenvironments. For example, MID positions with <20% εA excision are located near histone tails; sites −34, −37, −47, and −111 are closest to the base of an H2A tail, and position 119 is, similarly, near an H2B tail. Indeed, other glycosylases (27, 30) and BER enzymes (31) have shown variable amounts of inhibition due to histone tails. The proximity of histone tails to certain positions may interfere with AAG in similar ways to reduce activity.

We also used DNase I footprinting as a complementary approach to HRF to characterize the solution accessibility of nucleobases using an enzymatic probe of similar size to that of AAG (molecular weight of DNase I is 29 kDa versus 26 kDa for AAG) (Figures 4.11 and 4.12). The local maxima of accessibility within a helical
Figure 4.10 HRF denaturing PAGE gels to characterize nucleobase solution accessibilities in global εA-containing NCPs. The HRF reaction on the global εA-containing I strand NCPs of the (A) 5'-end and (B) 3'-end are shown in the respective HRF lanes. The HRF reaction on the global εA-containing J strand NCPs of the (C) 5'-end and (D) 3'-end are shown in the respective HRF lanes. The L lanes display a sequence ladder created using Maxam-Gilbert reactions using undamaged Widom 601 sequences.
turn are consistent with those observed by HRF (Figure 4.13).

The overall correlation between glycosylase activity and solution accessibility observed in these in vitro chromatin experiments is consistent with in vivo observations. A study in yeast mapped alkylation damage and its subsequent repair by the yeast homologue of AAG across the genome (32); the examination of 10000 strongly positioned NCPs in the genome revealed decreased repair at less solution-accessible sites. Another recent study reported that somatic mutational rates in ~3500 human tumors are enriched in sequences packaged by NCPs (33). Furthermore, a subset of tumor types had a relative increase in mutation rates where minor grooves faced the histone core, which is specifically attributed to decreased repair at less-accessible positions.

A reduction of AAG activity within chromatin may lead to this imbalance in DNA damage and repair that could cause mutagenicity and genetic instability described in vivo (32, 34). Indeed, DNA packaged in an NCP has been shown to be as susceptible to alkylating agents as unpackaged DNA (35). In this work, we have observed that AAG can remove εA lesions from HIGH positions, but has minimal activity on LOW positions. These results suggest that εA repair by AAG at LOW positions in vivo may require additional nuclear factors, such as histone variants, histone post-translational modifications, and/or chromatin remodelers to increase access to LOW-positioned lesions (36, 37).

Interestingly, these experiments with global εA lesion NCPs reveal exceptions to the correlation between solution accessibility and AAG activity. Specifically, there are instances of HIGH positions that are inefficiently excised by AAG. The first such position is −74, located near the dyad axis, where only 20% εA excision was observed. The dyad region has been shown to have different characteristics from the rest of the NCP, including less DNA-histone dynamics (11), a straighter conformation, and being
Figure 4.11 DNase I footprinting on 601I DUP and NCPs. The DNase I footprinting reaction on the (A) 5’-end and (B) 3’-ends. DNA substrates (2.5 pmol) were reacted with 0.002 U or 0.02 U of DNase I, as shown in E lanes. The N lane is an untreated DUP loaded for reference. L lanes are Maxam-Gilbert sequencing reactions, whose products (labeled on the left side of each gel) migrate about 1 nt faster than their analogous DNase I cleavage products.
Figure 4.12 DNase I footprinting on 601J DUP and NCPs. The DNase I footprinting reaction on the (A) 5' end and (B) 3' ends. DNA substrates (2.5 pmol) were reacted with 0.002 U or 0.02 U of DNase I, as shown in E lanes. The N lane is an untreated DUP loaded for reference. The L lanes are Maxam-Gilbert sequencing reactions, whose products (labeled on the left side of each gel) migrate about 1 nt faster than their analogous DNase I cleavage products.
Figure 4.13 Quantification of DNase I footprinting of the 60I strand (red crosses) and the 601J (yellow crosses) in NCPs is compared with AAG excision (green and blue bars) and HRF profile (gray area). DNase I activity and the HRF profile share the y-axis on the right.
underwound (9), any of which may contribute to suppressed AAG activity. Consistent with previous studies with restriction enzymes (10), there is also an overall decrease in DNase I cleavage within the dyad region (Figure 4.13). Indeed, we (13, 17, 29) and other researchers (28, 38, 39) have found that the repair of lesions in the ~20 bp of DNA centered at the dyad region is suppressed, relative to other regions of the NCP. *In vivo* observations have also shown these patterns of decreased BER within the dyad region of strongly positioned NCPs (32).

Sites -30, 96, and 97 are also HIGH positions that are inefficiently excised by AAG, but they are located outside of the inhibitory dyad region. In addition, all three positions are adjacent to sites where εA is more readily excised. For instance, AAG excised 40% of εA at position -30, but nearly quantitative conversion to product was observed at position -31. These two positions are potentially influenced by the unstructured tails of nearby H2B and/or H2A (Figure 4.14A). In addition, these two positions are close to superhelical location (SHL) 5, which is a location in nucleosomal DNA known for stretching and disrupted base stacking (25, 40). Specifically, the local differences in the dinucleotide step parameters of twist, roll, and slide between the consecutive positions could preferentially prevent AAG excision at position -30. In addition, DNase I activity in this region was not as robust as in other helical turns on this strand (Figure 4.13), which might suggest a local NCP structural feature that inhibits general enzymatic activity. Similarly, AAG excised 50% of εA at position 96, while only removing 20% εA at position 97. Reduced εA excision here may be due to the nearby H2B and/or H4 tail (Figure 4.14B). Alternatively, DNA stretching is also known to occur nearby between SHL ± 1 to SHL ± 2, to result in severe kinking, and could be a direct result of the very different roll parameters observed between positions 96 and 97 (25, 40), and thereby inhibit AAG activity here. It is also of note that both the HRF profile and DNase I activity in this region were high (Figure 4.13)
and could indicate that the local NCP architecture is specifically inhibiting AAG activity. Generally, across the sequence, the very extreme dinucleotide step parameters of twist, roll, and slide analyzed in 601 NCP structures,\textsuperscript{14,30} such as the low degree roll at SHL ± 4.5 (i.e., around positions 28 or 119), or minor groove-inward pressure point positions, do not appear to influence AAG activity beyond the associated lower-solution accessibility. All other consecutive run εA positions observed, such as 28 and 29; 58 through 60; 89 and 90; −56 and −57; −85 through −88; and −103 and −104, have εA excision correlated with relative solution accessibility, as expected. The differential εA excision by AAG in the described positions emphasizes the effects of subtle, local NCP microenvironments unique to each nucleobase position and highlights the complexities that DNA packaging into an NCP presents for a glycosylase. Instances of decreased AAG activity in HIGH positions create an imbalance between damage and repair that can contribute to mutagenic signatures (33), and describe potential hotspots in the packaged genome for persistent alkylation damage due to a deficiency in repair.

The ends of DNA in NCPs have been shown to spontaneously and transiently unwrap to allow for increased access to nucleobases (11). Furthermore, NCPs asymmetrically unwrap \textit{in vivo} (41) and \textit{in vitro} (42, 43), meaning that one end is more dynamic than the other. In our global εA NCPs, we observed increased AAG activity at the end known to unwrap preferentially in the 601 NCP (42, 43). From base pair 123 to the DNA end, there is increased repair, regardless of solution accessibility (Figure 4.8). Specifically, at least 60% excision of εA was observed at LOW positions, where similar solution-accessible positions elsewhere in the NCP, at positions such as 28, 29, −45, 48, 59, −67, −88, had, at most, 30% excision. MID positions on this end, such as 123, 126, −135, showed at least 80% εA excision, while MID positions on the other end, such as 19, −24, −34, had, at most, 20%. These results are consistent
Figure 4.14 Proximity of histone tails to HIGH positions with low AAG activity. Zoom-in views of NCP crystal structure showing proximity of histone tails that may be within the potential area of effect on HIGH positions with decreased AAG activity. (A) Positions −30 and −31 (shown as gray sticks with a surface overlay) are in proximity to the nearby tails of H2A (yellow) and/or H2B (red). (B) Positions 96 and 97 (shown as gray sticks with a surface overlay) are similarly near the tails of H4 (green) and H2B (red). Only the histones and lesion-containing strand are shown for simplicity. The histone tails are shown with a surface overlay for emphasis. A merged NCP crystal structure of the Widom 601 duplex (PDB code 3lz0) and histone octamer with tails (PDB 1kx5) was created for these images using PyMOL.
with *in vivo* observations of higher levels of repair asymmetric to the dyad (32).

### 4.5 Conclusion

In summary, we determined the repair profile and various kinetic parameters in global DUP and NCPs to observe εA excision by AAG activity at 49 distinct geometric positions. AAG activity across the DUP had maximal amounts of excision, but varying amounts of activity across the NCP architecture. We show that the presence of the histone octamer, and therefore the packaging of DNA into chromatin, modulates repair by AAG. Although AAG’s monophasic kinetic behavior remained consistent across the NCP, we observed interesting relationships between the amounts of εA excision and NCP architecture. More specifically, we observed a general correlation between AAG activity and solution accessibility in the NCPs. However, we observed a few exceptions to this correlation, including HIGH positions with low repair likely due to distinct NCP microenvironments that suppress AAG activity, and increased εA removal at the DNA end that asymmetrically unwraps. Our kinetic characterizations of AAG excision of εA in global NCPs have revealed interesting microenvironments that inhibit repair that could not be appreciated previously in site-specific lesion NCPs. Future studies using this global NCP system could provide a more comprehensive understanding of the nuanced relationships between NCP architecture and DNA repair in cells.
4.6 Supporting information

DNA sequences used in the present work

I and J strand designations are based on the crystal structure of Vasudevan, et al. for the Widom 601 NCP (25).

I strand

5'- ATC AGA ATC CCG GTG CCG AGG CCG CTC AAT TGG TCG TAG ACA 
GCT CTA GCA CCG CTT AAA CGC ACG TAC GCG CTG TCC CCC GCG TTT 
TAA CCG CCA AGG GGA TTA CTC CCT AGT CTC CAG GCA CGT GTC AGA 
TAT ATA CAT CGA T

J strand

5'- ATC GAT GTA TAT ATC TGA CAC GTG CCT GGA GAC TAG GGA GTA 
ATC CCC TTG GCG GTT AAA ACG CGG GGG ACA GCG CGT ACG TGC 
GTT TAA GCG GTG CTA GAG CTG TCT AGT CTC CAG GCA CGT GTC AGA 
TAT ATA CAT CGA T

i.1 (35 mer)

5'- ATC AGA ATC CCG GTG CCG AGG CCG CTC AAT TGG TC

i.2 (45 mer)

5'- GTA GAC AGC TCT AGC ACC GCT TAA ACG CAC GTA CGC GCT GTC CCC
i.3 (65 mer)

5’- CGC GTT TTA ACC GCC AAG GGG ATT ACT CCC TAG TCT CCA GGC
ACG TGT CAG ATA TAT ACA TCG AT

i.s12 (32 mer)

5’- GTG CTA GAG CTG TCT ACG ACC AAT TGA GCG GC

i.s23 (26 mer)

5’- GGC GGT TAA AAC GCG GGG GAC AGC GC

j.1 (45 mer)

5’- ATC GAT GTA TAT ATC TGA CAC GTG CCT GGA GAC TAG GGA GTA
ATC

j.2 (65 mer)

5’- CCC TTG GCG GTT AAA ACG CGG GGG ACA GCG CGT ACG TGC GTT
TAA GCG GTG CTA GAG CTG TCT AC

j.3 (35 mer)

5’- GAC CAA TTG AGC GGC CTC GGC ACC GGG ATT CTG AT

j.s12 (30 mer)

5’- TTT AAC CGC CAA GGG GAT TAC TCC CTA GTC
j.s23 (32 mer)

5'- GCC GCT CAA TTG GTC GTA GAC AGC TCT AGC AC

Internal reference oligonucleotide 23 mer

5'- ATC AGA ATC CCG GTG CCG AGG CC

Internal reference oligonucleotide 53 mer

5'- ATC GAT GTA TAT ATC TGA CAC GTG CCT GGA GAC TAG GGA GTA ATC CCC TTG GC

Internal reference oligonucleotide 92 mer

5'- ATC AGA ATC CCG GTG CCG AGG CCG CTC AAT TGG TCG TAG ACA GCT CTA GCA CCG CTT AAA CGC ACG TAC GCG CTG TCC CCC GCG TTT TAA CC

Figure 4.15 Ligation strategy for complementary 145 mer oligonucleotides lacking damage. Each 145 mer is assembled by the ligation of three component oligonucleotides. Components 2 and 3 are 5'-phosphorylated and combined with component 1 and two scaffold oligonucleotides s12 and s23 for annealing. Annealed strands are subsequently ligated to create the full-length strands. When global εA was in the I strand, it was annealed to the non-lesion-containing J strand. Similarly, when global εA was in the J strand, it was annealed to the non-lesion-containing I strand.
4.7 References


Chapter 5

Conclusions and Future Perspectives

5.1 Conclusions

Genomic DNA is constantly assaulted by damaging agents derived from both cellular metabolism and the environment. Exposure to these damaging agents can result in a variety of chemical modifications to DNA nucleobases. The base excision repair pathway (BER) is responsible for the repair of non-bulky modified nucleobases (nucleobase lesions), which is initiated by a DNA glycosylase through lesion recognition and excision. In this work, we have examined the initiation of BER by different glycosylases within the context of packaged DNA. The influence of several factors on glycosylase activity has been investigated, for instance, histone post-translational modification (PTM) and histone variant incorporation. Using nucleosome core particle (NCP), the primary repeating unit of chromatin, as a model system, we have studied four glycosylases, human oxoguanine DNA glycosylase (OGG1), uracil DNA glycosylase (UDG), single-strand selective monofunctional uracil DNA glycosylase (SMUG1), and alkyladenine DNA glycosylase (AAG), repairing 8-oxo-7,8-dihydroguanine (8oxoG), uracil (U), and 1,N6-ethenoadenine ($\varepsilon$A), respectively, each of which are representative lesions caused by oxidation, deamination and alkylation. The global fingerprint approach and kinetic experiments have provided us a systematic and comprehensive platform to define BER in packaged DNA and elucidate to which extent those cellular factors can modulate the initiation of BER by glycosylases.

We started with creating homogeneous NCP using the Widom 601 nucleosome positioning sequence and canonical histone proteins that were recombinantly expressed and purified. The lesion was positioned off the dyad axis (~20 bp from the DNA end) and in one of three rotational positions: out toward solution (OUT), in toward the histone core (IN), or somewhere in between (MID). In contrast to the previous observation that OGG1 activity is completely inhibited for the lesions on the
dyad (1), OGG1 can excise 8oxoG at all three rotational positions off the dyad without assistance from other ancillary cellular components. We attribute this increase in OGG1 repair to the transient unwrapping of nucleosomal DNA toward the DNA end, which allows for the exposure of lesion sites, especially those that are structurally occluded, to the enzyme. We also noticed a disparity between lesion solution accessibility and OGG1 activity: the MID position exhibited the highest level of 8oxoG repair while the OUT position showed the lowest, indicating the influence of other regulatory elements on determining glycosylase reactivity. We therefore examined structural features of these off-dyad lesion positions being guided by a merged crystal structure model of an NCP that contains Widom 601 DNA and histone tails. In this model we observed that the off-dyad lesions, especially the OUT position, are near an H2B tail that protrudes between the superhelices. However, depletion of H2B tails failed to restore OGG1 activity; in contrast, chemical acetylation of H2B proteins lead to a significant increase in 8oxoG repair at the OUT position, suggesting that the neutralization of lysine residues within H2B and the resulting relaxation of the NCP structure contribute to this phenomenon. We also performed single-turnover kinetic assay to measure the observed rate of OGG1 excising 8oxoG positioned off the dyad axis. While the data for free duplex DNA control fit well to a single exponential model, double exponential models were required to fit the data for all lesion positions in an NCP. We attribute the biphasic kinetics to the existence of two distinct NCP populations. The fast phase exhibited comparable rates to that on free duplex, which we attribute to a configuration that can be processed readily by the glycosylase. The slower phase was attributed to a population of NCP that require a rate-limiting conformational change to adopt a structure permissive for repair. This conformational change is consistent with the transient unwrapping of DNA to expose the off-dyad lesion sites, and is not observed in experiments with lesions positioned
on the dyad (1). Overall, these studies using NCP containing a single lesion at a specific location highlight the complexity of the NCP environment for glycosylases. Multiple regulatory elements, such as the transient unwrapping of nucleosomal DNA and the presence of histone tails, may work synergistically to modulate the initiation of BER in the context of packaged DNA.

An important aspect we learned from the experiments using NCP containing a site-specific lesion, as outlined above, is that though they are informative in demonstrating glycosylase performance at a defined position, there is no lesion site that can be representative of repair across all sites within an NCP. In fact, factors governing initiation of BER in NCP at a site are more nuanced and complex than previously appreciated and are strongly guided by local NCP dynamics. To gain an overall picture of the repair profile of a glycosylase, we developed an approach to globally incorporate a lesion into NCP by chemical synthesis such that multiple lesion sites can be examined simultaneously. With this approach, we prepared a population of NCPs that contains U with a variety of translational and rotational positions, to investigate the impact of histone H2A variant incorporation on U excision by UDG and SMUG1. As introduced previously, substituting histone variants for their canonical counterparts can profoundly alter chromatin structure, thereby impacting multiple biological processes including DNA repair. The two H2A variants studied in this work, H2A.Z and macroH2A, were separately incorporated into NCP, for the replacement of canonical H2A. We found that while SMUG1 activity is substantially suppressed along the entire sequence in canonical NCP, U excision by UDG from canonical NCP is largely dominated by solution accessibility of U: sites that are highly solution-accessible, as determined by hydroxyl radical footprinting, exhibited high level of U repair, and sites that are solution-inaccessible exhibited low level of U repair. Interestingly, the region
flanking the dyad axis is a notable exception to solution accessibility dictating glycosylase activity; excision of U by UDG is significantly inhibited regardless of solution accessibility in this region. In contrast, the U sites that showed limited UDG activity in canonical NCP are more readily excised in macroH2A and H2A.Z NCPs, reflecting the ability of these variants to facilitate excision at sites that are otherwise poorly repaired. We also observe that U with the largest increase in excision in variant NCPs are clustered in regions with differential structural features between the variants and canonical H2A. Within 35-40 bp from the DNA end in macroH2A NCP, the activities of both UDG and SMUG1 are comparable to that observed on a free duplex control. We showed that this high level of activity results from two distinct species within the macroH2A NCP ensemble: octasomes and hexasomes. These observations reveal potential functions for H2A variants in promoting BER and preventing mutagenesis within the context of chromatin.

We then expanded the application of the global approach by combining it with kinetic assay, to evaluate the global profile of AAG repair in packaged DNA. In this work, we assembled a global population of NCPs in which collectively each A is replaced with εA. While each NCP contains no more than one εA lesion, the total population contains εA in 49 distinct geometric positions. Similar to the previous observations for UDG excising U, AAG reactivity is largely correlated with solution accessibility of εA across the NCP. Nevertheless, we did identify some highly solution-accessible lesions that are not repaired well, presumably due to the interference of histone tails nearby. We also noticed an increase in repair within the region toward the DNA end that is known to possess higher hydrodynamics and preferentially unwrap (2–4). In the kinetic assay, AAG excising εA displayed monophasic behavior across the NCP, but with varying repair level, suggesting that there are subpopulations of NCPs that are conformationally competent for AAG activity while others are not. We
attribute the lack of biphasic kinetics of AAG to its slower activity, compared to that of OGG1, so that the slower kinetic phase of AAG in which the conformational change may be the rate-determining step cannot be resolved on the time scale in the current experiments. Taken together, these observations obtained from the global approach reveal the structural nuances of NCP environment and provide molecular-level insight into the relationship between glycosylase repair and NCP architecture.

5.2 Future Perspectives

Unlike the understanding that has been gained for DSB repair and NER, progress in elucidating the behavior of BER enzymes has been slower. This slower progress results part from the technical difficulty of creating global and/or site-specific damage in higher order chromatin models and from the lability of certain nucleobase lesions, which are challenging to manipulate both in vivo and in vitro. On the other hand, the lack of strong chromatin signals, for instance, phosphorylation of H2A.X when DSB takes place (5), makes it hard to monitor the occurrence of BER in cells. All of these considerations call for advanced biochemical techniques, synthetic chemistry, and experimental design.

While the influence of some factors on glycosylase activity, such as lesion translational and rotational positioning, is well-understood, the examination of other factors has just begun, including histone PTMs, histone variant incorporation, and chromatin remodeling. It is tempting to speculate that BER might be coupled with other cellular events, such as transcription. This is supported by the observation that Endonuclease VIII-like glycosylase 2 (NEIL2) preferentially repairs 5-hydroxyuracil (5hU) in transcribed genes (6). Moreover, AAG has been observed to localize at chromatin and engage in a complex with actively transcribing RNA polymerase II
(7), providing direct evidence of BER associating with transcription elongation. The switch of histone chaperone FACT (facilitates chromatin transcription) from binding with transcription factors to binding with repair enzymes upon 8oxoG formation (8) also suggests the coordination between transcription and BER. Investigating transcription-associated BER will provide insights into how BER is regulated by histone PTMs, variant incorporation, and chromatin remodeling, which have been shown to be related to transcription (9–12). On the other hand, it has been reported that OGG1 activity in excising 8oxoG is increased through the interaction between OGG1 and UV-DDB (ultraviolet light-damaged DNA binding protein) (13), a key protein factor in human NER, revealing potential coordination between BER and NER.

It is also of interest to explore the initiation of BER in specific sections of chromosomes, such as the telomeres where DNA adopts a different higher order architecture (14–16) as compared to nucleosomal DNA. Telomeric DNA contains tandem repeats of 5′-TTAGGG-3′ (17) and are subject to oxidative guanine damage (18). In fact, it has been shown that 8oxoG formation in telomeric DNA destabilizes the G-quadruplex folding, alters telomerase extension capability, and ultimately affect telomere homeostasis (19, 20). Given the distinct biological roles of telomere, studying BER in the telomere area may lead us to a better understanding of the link of BER to disease, cancer and aging.
5.3 References


