

From Nucleotide to Nucleosome:
A Kinetic and Calorimetric Study of DNA

By Ji Huang

M.S. Nanjing University 2010

B.S. Nanjing University 2007

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 2016

© Copyright 2016 Ji Huang

This dissertation by Ji Huang is accepted in its present form by the
Department of Chemistry as satisfying the dissertation requirement for the degree of
Doctor of Philosophy

Recommended to the Graduate Council

Date _____
Sarah Delaney, Ph.D. Advisor

Date _____
Jason Sello, Ph.D. Reader

Date _____
Christopher Seto, Ph.D. Reader

Approved by the Graduate Council

Date _____
Peter M. Weber, Ph.D.
Dean of the Graduate School

Curriculum Vitae

Ji Huang

Education

Brown University, Providence, Rhode Island
Ph.D., Chemistry, expected May 2016

Nanjing University, Nanjing, China
M.S., Chemistry, May 2010

Nanjing University, Nanjing, China
B.S., Chemistry, May 2007

Research Experience

Department of Chemistry, Brown University, Providence, Rhode Island
Graduate Student with Dr. Sarah Delaney, 2011-present

- Investigate the thermodynamic properties of DNA containing various number of trinucleotide repeats from DNA only system to the nucleosome system by differential scanning calorimetry.
- Elucidate by kinetic techniques of how polymerase discriminates the hyperoxidized nucleotide from the canonical nucleotide pool to prevent DNA mutation.

Department of Chemistry, Nanjing University, Nanjing, China
Master Student with Dr. Li-Min Zheng & Dr. Lin Guo, 2007-2010

- Apply bacterial magnetic particles for rapid enrichment of phosphopeptides.
- Develop methods for surface modification of bacterial magnetic nanoparticles as drug carriers for therapeutic applications.

Publications

- J. Huang, S. Delaney (2016) Unique length-dependent biophysical properties of repetitive DNA. *J. Phys. Chem. B*, accepted, DOI: 10.1021/acs.jpcc.6b00927.
- J. Huang, C. J. Yennie, S. Delaney (2015) Klenow fragment discriminates against the incorporation of the hyperoxidized dGTP lesion spiroiminodihydroantoin into DNA. *Chem. Res. Toxicol.* 28, 2325-2333
- M. E. Tarantino, K. Bilotti, J. Huang, S. Delaney (2015) Rate-determining step of flap endonuclease 1 (FEN1) reflects a kinetic bias against long flaps and trinucleotide repeat sequences. *J. Biol. Chem.* 290, 21154-21162
- L. Guo, J. Huang, L. M. Zheng (2012) Bifunctional bacterial magnetic nanoparticles for tumor targeting. *Nanoscale* 4, 879-884
- L. Guo, J. Huang, L. M. Zheng (2011) Control generating of bacterial magnetic nanoparticle-doxorubicin conjugates by poly-L-glutamic acid surface modification. *Nanotechnology* 22, 175102

- J. Huang, L. Guo, L. M. Zheng (2010) Rapid enrichment and determination of phosphopeptides using bacterial magnetic particles *via* both strong and weak interactions. *Analyst* 135, 559-563
- L. Guo, J. Huang, L. M. Zheng (2010) Efficient conjugation doxorubicin to bacterial magnetic nanoparticles via a triplex hands coupling reagent. *J. Nanosci. Nanotechnol.* 10, 6514-6519
- Z. Y. Li, M. Y. Ten, J. J. Ma, J. Huang, L. Y. Wang, Y. Pan (2009) Unprecedented synthesis of chiral calix[4](aza)crowns and its potent encapsulating methanol. *Sci. China Ser. B-Chem.* 52, 497-504
- L. Guo, J. Huang, X. Zhang, Y. Li, L. M. Zheng (2008) Baterial magnetic nanoparticles as drug carriers. *J. Mater. Chem.* 18, 5993-5997

Poster Presentations

- J. Huang, S. Delaney (2016) Klenow fragment discriminates against the incorporation of the hyperoxidized dGTP lesion spiroiminodihydantoin into DNA. *Gordon Research Conference on DNA Damage, Mutation & Cancer*, Ventura, CA, March 12-18
- J. Huang, S. Delaney (2014) Calorimetric study of nucleosome and its subunits. *13th International Workshop on Radiation Damage to DNA*, MIT Cambridge, MA, June 14-18
- J. Huang, S. Delaney (2014) Calorimetric study of nucleosome. *247th ACS National Meeting*, Dallas, TX, March 16-20
- J. Huang, S. Delaney (2013) Thermodynamic and kinetic studies of trinucleotide repeat (TNR) DNA. *Journal of Biomolecular Structure and Dynamics*, 18th Conversation, Albany, NY, June 9-13
- J. Huang, S. Delaney (2012) Calorimetric and kinetic studies of trinucleotide repeat (TNR) DNA *Gordon Research Conference on Mutagenesis*, New Port, RI, August 19-24

Awards and Fellowships

- People's Scholarship, Nanjing University 2003-2005
- Graduate Student Scholarship, Nanjing University 2008-2010
- Brown University Graduate Fellowship, Brown University, 2010

Teaching Experience

- Teaching assistant of general chemistry lab at Nanjing University 2007-2008
- Teaching assistant of general chemistry lab at Brown University 2010-2011
- Teaching assistant of organic chemistry lab at Brown University 2015

Preface and Acknowledgements

To my advisor, labmates, family, and friends.

Table of Contents

| | |
|--|-----|
| Signature Page..... | iii |
| Curriculum Vitae..... | iv |
| Preface and Acknowledgements..... | vi |
| Table of Contents..... | vii |
| List of Tables..... | xi |
| List of Figures..... | xii |
| Chapter 1. Introduction | |
| 1.1 Introduction to DNA and DNA Damage..... | 2 |
| 1.2 Guanine Oxidation..... | 3 |
| 1.3 DNA Damage from Nucleotide Pool..... | 6 |
| 1.4 Repair of 8-oxoG and Sp Lesions..... | 7 |
| 1.5 Chemical and Kinetic Mechanism of Polymerases..... | 8 |
| 1.6 Techniques for Studying Nucleotide Incorporation Kinetics..... | 9 |
| 1.7 Introduction to Trinucleotide Repeat (TNR) DNA..... | 12 |
| 1.8 TNR Expansion Mechanisms..... | 14 |
| 1.9 Nucleosome—Basic Unit for DNA Packaging..... | 18 |
| 1.10 Introduction to Differential Scanning Calorimetry (DSC)..... | 19 |
| 1.11 Concluding Remarks and Gap in Knowledge..... | 22 |
| 1.12 References..... | 23 |

Chapter 2. Klenow Fragment Discriminates Against the Incorporation of the Hyperoxidized dGTP Lesion Spiroiminodihydantoin into DNA

| | |
|---|----|
| 2.1 Abstract..... | 38 |
| 2.2 Introduction..... | 39 |
| 2.3 Experimental Procedure..... | 41 |
| 2.3.1 Materials..... | 41 |
| 2.3.2 Oligonucleotide Synthesis and Purification..... | 41 |
| 2.3.3 Synthesis, Purification, and Absolute Configuration of Spiroiminodihydantoin-2'-deoxynucleoside-5'-triphosphate..... | 41 |
| 2.3.4 Preparation of DNA Primer/Template Assemblies..... | 45 |
| 2.3.5 Qualitative Single Nucleotide Triphosphate Incorporation Reactions..... | 45 |
| 2.3.6 Primer Extension after the Incorporation of Sp Nucleotide Triphosphate..... | 46 |
| 2.3.7 Single-Turnover Kinetics Reactions..... | 46 |
| 2.3.8 Data Analysis..... | 46 |
| 2.4 Results and Discussion..... | 47 |
| 2.4.1 Qualitative Analysis of Incorporation of Spiroiminodihydantoin Nucleotide Triphosphate by Klenow Fragment..... | 47 |
| 2.4.2 Kinetic Analyses of Nucleotide Triphosphate Incorporation..... | 51 |
| 2.4.3 Single-Turnover Kinetic Analysis of dGTP and 8-oxodGTP Incorporation..... | 52 |
| 2.4.4 Single-Turnover Kinetic Analysis of dSpTP Incorporation..... | 55 |
| 2.4.5 Comparing the Kinetic Parameters for dGTP, 8-oxodGTP, (S)-dSpTP, and | |

| | |
|--------------------------------------|----|
| (R)-dSpTP..... | 58 |
| 2.4.6 Biological Considerations..... | 60 |
| 2.5 Conclusion..... | 61 |
| 2.6 Reference..... | 61 |

Chapter 3. Unique Length-Dependent Biophysical Properties of Repetitive DNA

| | |
|--|----|
| 3.1 Abstract..... | 73 |
| 3.2 Introduction..... | 74 |
| 3.3 Experimental Procedures..... | 76 |
| 3.3.1 Oligonucleotide Synthesis and Purification..... | 76 |
| 3.3.2 Differential Scanning Calorimetry (DSC) Analysis..... | 77 |
| 3.3.3 Native Polyacrylamide Gel Electrophoresis..... | 79 |
| 3.3.4 Kinetics of Hairpin to Duplex Conversion..... | 79 |
| 3.4 Results and Discussion..... | 80 |
| 3.4.1 DSC Characterization of (CAG) _n and (CTG) _n Stem-Loop Hairpins..... | 81 |
| 3.4.2 Electrophoretic Mobility of (CAG) _n and (CTG) _n by Native PAGE..... | 82 |
| 3.4.3 Determining the Nature of Homoduplex Melting..... | 85 |
| 3.4.4 Analysis of DSC Thermograms for (CAG) _n and (CTG) _n Stem-Loop Hairpins..... | 87 |
| 3.4.5 Thermodynamics of Stem-Loop Hairpin to Duplex Conversion..... | 91 |
| 3.4.6 Kinetic Analysis of Stem-Loop Hairpin to Duplex Conversion..... | 94 |
| 3.5 Conclusion..... | 98 |

| | |
|---------------------|----|
| 3.6 References..... | 99 |
|---------------------|----|

Chapter 4. Differential Scanning Calorimetry Study of Nucleosome and its Subunits

| | |
|--|-----|
| 4.1 Abstract..... | 104 |
| 4.2 Introduction..... | 105 |
| 4.3 Experimental Procedure..... | 107 |
| 4.3.1 Preparation of Nucleosomal DNA..... | 107 |
| 4.3.2 Histone Expression and Purification..... | 109 |
| 4.3.3 Histone H2A/H2B Dimer, (H3-H4) ₂ Tetramer and Histone Octamer Assembly..... | 110 |
| 4.3.4 Nucleosome Reconstitution and Purification..... | 110 |
| 4.3.5 DSC Characterization of Nucleosome..... | 111 |
| 4.4 Results and Discussion..... | 112 |
| 4.4.1 Preparation of Nucleosomal DNA..... | 112 |
| 4.4.2 DSC Characterization of Nucleosomal DNA..... | 113 |
| 4.4.3 DSC Characterization of Histone H2A/H2B Dimer, (H3-H4) ₂ Tetramer and Octamer..... | 118 |
| 4.4.4 DSC Characterization of Nucleosome..... | 121 |
| 4.5 Conclusion..... | 126 |
| 4.6 References..... | 126 |

List of Tables

| | |
|---|-----|
| Table 1.1 Representative Disease Caused by TNR Expansion..... | 13 |
| Table 2.1 Kinetic Parameters for the Incorporation of dNTPs Opposite Template dC..... | 59 |
| Table 3.1 DSC-Derived Melting Temperatures and Thermodynamic Parameters for (CTG) _n and (CAG) _n Stem-Loop Hairpins Under Physiological Relevant Salt Condition..... | 85 |
| Table 3.2 DSC-Derived Melting Temperatures and Thermodynamic Parameters for (CTG) _n and (CAG) _n Stem-Loop Hairpins under low salt conditions..... | 90 |
| Table 3.3 DSC-Derived Melting Temperatures and Thermodynamic Parameters for (CAG) _n /(CTG) _n Duplexes under physiological relevant salt condition..... | 93 |
| Table 3.4 DSC-Derived Melting Temperatures and Thermodynamic Parameters for (CAG) _n /(CTG) _n Duplexes under low salt conditions..... | 94 |
| Table 3.5 Rate and Activation Enthalpy for Stem-Loop Hairpin to Duplex Conversion at 37 °C..... | 95 |
| Table 4.1 DNA Sequences Used in this Study..... | 108 |

List of Figures

| | |
|--|----|
| Figure 1.1 DNA Double Helix Structure, Base Pair Pattern and Nomenclature..... | 2 |
| Figure 1.2 Base Pair of G to C and 8-oxoG to C or A..... | 4 |
| Figure 1.3 Oxidation Pathway of Generating 8-oxoG and Sp from G..... | 5 |
| Figure 1.4 Kinetic Scheme of DNA Polymerization..... | 9 |
| Figure 1.5 Techniques used to Study Kinetic Mechanism of Polymerase..... | 11 |
| Figure 1.6 Hairpin Structure Formed by (CNG) _n Repeat..... | 15 |
| Figure 1.7 Representative Replication- and Repair-Dependent Models for TNR Expansion..... | 16 |
| Figure 1.8 Crystal Structure of Nucleosome..... | 18 |
| Figure 1.9 Simplified Working Principal of DSC..... | 20 |
| Figure 1.10 Representative DSC Thermogram and Data Analysis..... | 21 |
| Figure 2.1 Structures of dG, 8-oxodG, (<i>S</i>)-dSp, and (<i>R</i>)-dSp..... | 39 |
| Figure 2.2 HPLC Purification and Characterization of dSpTP Diastereomers..... | 43 |
| Figure 2.3 ESI-MS Analysis for dSp Diastereomers in the Nucleotide Forms or Nucleoside Forms..... | 44 |
| Figure 2.4 Qualitative Assessment of dSpTP Diastereomers Incorporation..... | 47 |
| Figure 2.5 Denaturing PAGE Showing the Migration Difference of Single Nucleotide Incorporation Product of Adding either dGTP or (<i>S</i>)-dSpTP..... | 48 |
| Figure 2.6 Kinetic Analysis of dGTP and 8-oxodGTP Incorporation by KF ⁻ | 53 |
| Figure 2.7 3-D Confidence Contours for the Rate Constants Determined by KinTek Global Explorer for dGTP and 8-oxodGTP..... | 55 |
| Figure 2.8 Kinetic Analysis of (<i>S</i>)-dSpTP and (<i>R</i>)-dSpTP Incorporation by KF ⁻ | 57 |

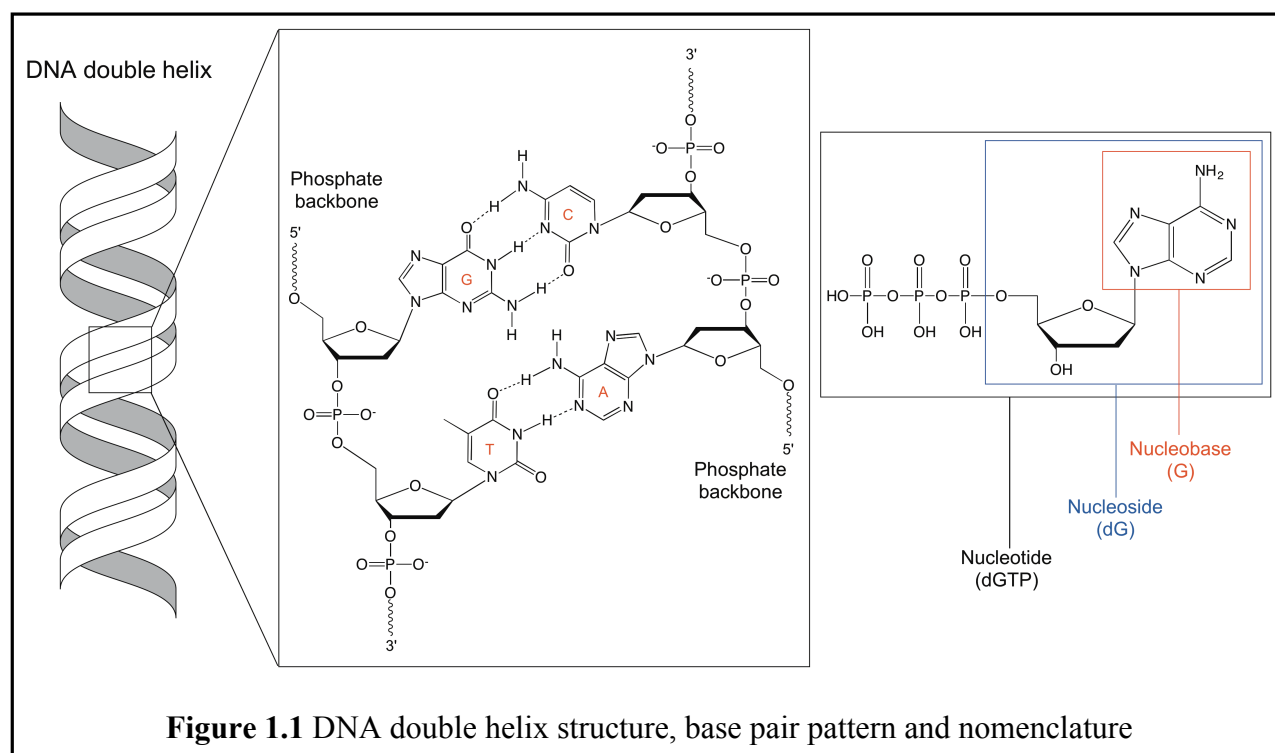
| | |
|---|----|
| Figure 2.9 Analysis of dGTP, 8-oxodGTP, (<i>S</i>)-dSpTP, and (<i>R</i>)-dSpTP Kinetic Results Using Michaelis-Menten Techniques..... | 59 |
| Figure 3.1 Representative Example of Baseline Line Creation for DSC Thermograms with Overlapped Transitions..... | 78 |
| Figure 3.2 Representative Example of Peak Deconvolution for DSC Thermograms with Overlapped Transitions..... | 78 |
| Figure 3.3 DSC Thermograms for (CTG) _n and (CAG) _n Sequences under Physiological Relevant Salt Condition where n = 6-14..... | 82 |
| Figure 3.4 Native PAGE Gel for (CTG) _n and (CAG) _n Sequences (n = 6-14) under Physiological Relevant Salt Condition..... | 83 |
| Figure 3.5 Native PAGE Characterization of (CTG) ₇ under Various Oligonucleotide Concentration and Buffer Conditions..... | 84 |
| Figure 3.6 Determination the Nature of Homoduplex Transition by Oligonucleotide Concentration Dependent Experiment..... | 86 |
| Figure 3.7 Enthalpy Change (ΔH) for Stem-Loop Hairpin (CTG) _n or (CAG) _n and Duplex (CTG) _n /(CAG) _n (n = 6-14) to Unstructured Single-Stranded DNA Transition..... | 88 |
| Figure 3.8 Native PAGE Gel for (CTG) _n and (CAG) _n Sequences (n = 6-14) under Low Salt Condition..... | 89 |
| Figure 3.9 Native PAGE Gel for (CTG) _n Sequences (n = 6-14) under Low Salt Condition..... | 89 |
| Figure 3.10 Native PAGE Gel for (CAG) _n Sequences (n = 6-14) under Low Salt Condition..... | 90 |
| Figure 3.11 DSC Thermograms of (CTG) _n /(CAG) _n (n = 6-14) duplex..... | 92 |
| Figure 3.12 Thermodynamic Parameters for Conversion of Stem-Loop Hairpins to Duplex Calculated via a Thermodynamic Cycle..... | 94 |

| | |
|--|-----|
| Figure 3.13 Kinetic Characterization of Stem-Loop Hairpin (CTG) _n and (CAG) _n to Duplex (CTG) _n /(CAG) _n (n = 6-14, 25, 30) Conversion..... | 96 |
| Figure 3.14 Determining the ΔH^\ddagger by Eyring Equation for (CTG) _n and (CAG) _n Stem-Loop Hairpin to (CAG) _n /(CTG) _n Duplex Conversion | 97 |
| Figure 3.15 Length Dependence of ΔH for the Conversion of Stem-Loop Hairpins to Canonical Duplex..... | 98 |
| Figure 4.1 Nucleosome Structure with Linker DNA and Linker Histone H1..... | 105 |
| Figure 4.2 DNA Constructs and Characterization..... | 112 |
| Figure 4.3 DSC characterization for S1 and 601 DNA and Their Deconvolution..... | 114 |
| Figure 4.4 DSC Characterization of Length Effect of S1 DNA on the Unfolding Progress..... | 115 |
| Figure 4.5 DSC Thermograms of S1, S1-10CAG, S1-20CAG and S1-30CAG DNA..... | 115 |
| Figure 4.6 The Effect of Salt Concentration on the Thermal Denaturation of Different DNA Sequences..... | 117 |
| Figure 4.7 The Effect of Salt Concentration on the Thermal Denaturation of H2A/H2B Dimer and (H3-H4) ₂ Tetramer..... | 118 |
| Figure 4.8 DSC Characterization of Thermal Denaturation of Histone Octamer under Different Salt Concentration..... | 119 |
| Figure 4.9 HPLC Analysis of DNA and Nucleosome (NCP)..... | 122 |
| Figure 4.10 DSC Characterization of S1 and 601 Nucleosome (NCP)..... | 122 |
| Figure 4.11 DSC Characterization of S1 Series Nucleosome (NCP) with 0, 10, 20 and 30 CAG Repeat..... | 123 |

Chapter 1: Introduction

1.1 Introduction to DNA and DNA Damage

DNA (deoxyribonucleic acid) was first discovered in 1869 by Swiss chemist Friedrich Miescher, but it was not until 1953 that the double helical structure (Figure 1.1) was solved by James Watson and Francis Crick.¹ In a subsequent paper by Francis Crick, he wrote “DNA is such an important molecule that is almost impossible to learn too much about it”.² Indeed, 60 years after the discovery of its structure, there is continuing interest in understanding the chemistry and biology of DNA. The DNA molecule stores genetic information in humans and almost all other organisms and it is essential to maintain its integrity for life development.



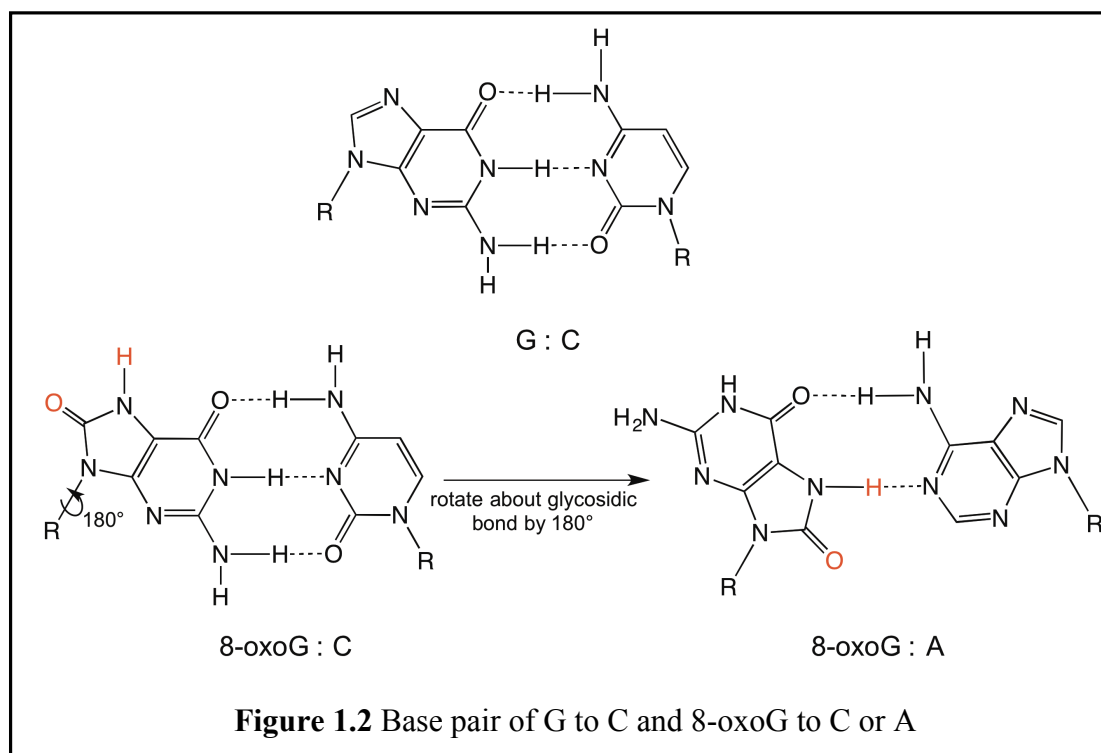
DNA is highly vulnerable to chemical modification and its integrity is endangered from several possible ways from the DNA backbone to the nucleobases. First, it is inevitable that DNA will undergo spontaneous hydrolysis but usually the half-lives are very long.³ Hydrolysis of the phosphodiester bonds on DNA backbone will cause single strand breaks; hydrolytic deamination of exocyclic amines on nucleobases will generate several products, with the most well-known

cytosine to uracil deamination; hydrolysis of the glycosidic bonds connecting nucleobases and sugars will generate abasic sites. Second, nucleobases in DNA will be modified by endogenous oxidizing agents from cellular respiration and inflammatory response.^{4,5} Reactive oxygen species (ROS) are a major component of the endogenous oxidizing agents, such as hydrogen peroxide (H_2O_2), superoxide ($\text{O}_2^{\bullet-}$), hydroxyl radical (OH^\bullet), singlet oxygen ($^1\text{O}_2$) and peroxynitrite (ONOO^-). Third, exogenous DNA oxidizing agents can also impair the integrity of DNA, and they include chemical compounds and high-energy electromagnetic radiation such as X-rays, ultraviolet light.⁶⁻⁹

1.2 Guanine Oxidation

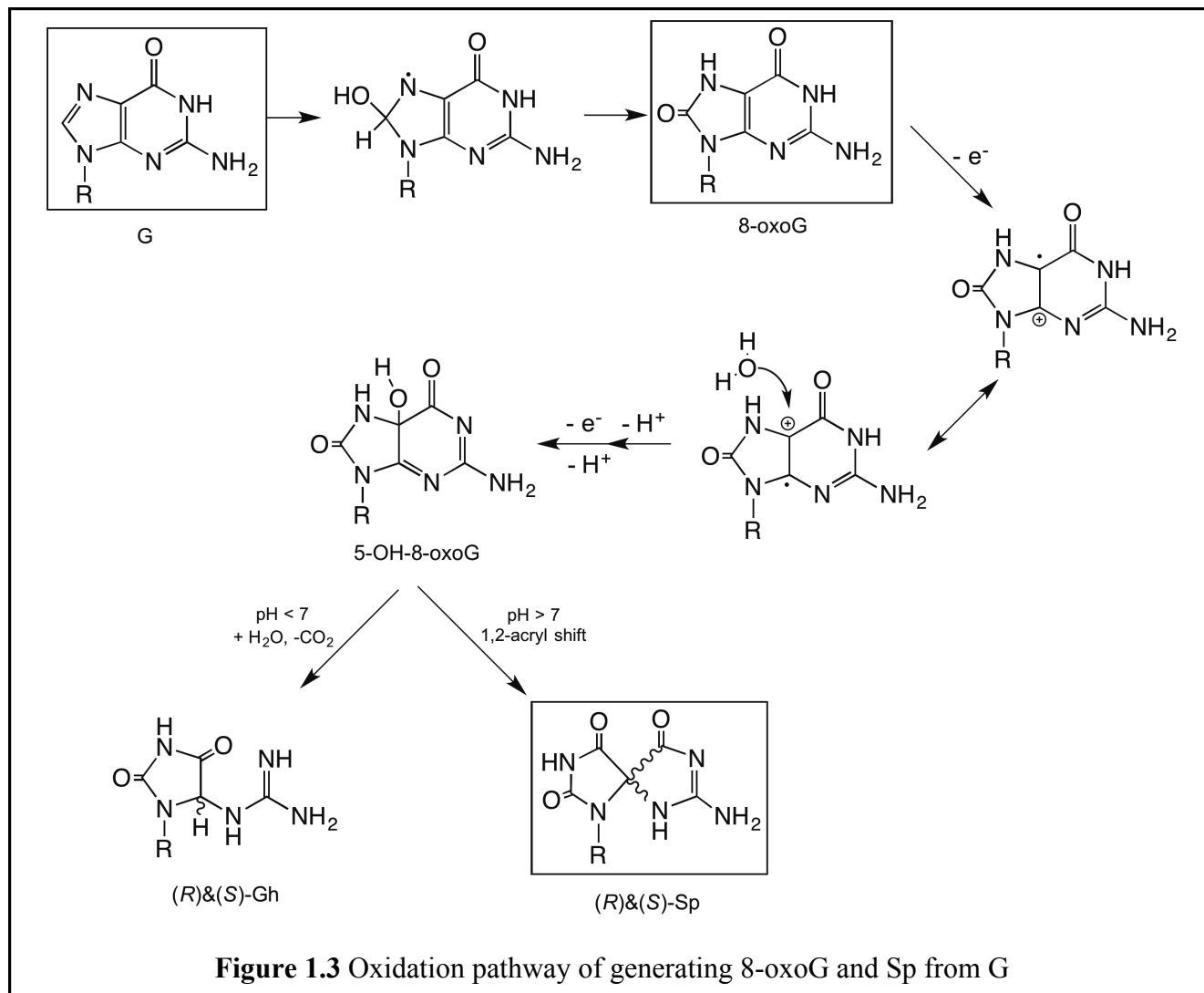
The one-electron reduction potentials of nucleosides in aqueous solution are $E_0 = 1.3, 1.4, 1.6$ and 1.7 V vs. NHE for dG, dA, dC and dT respectively, making guanine the most susceptible to oxidation among all four nucleobases.¹⁰ The most well-studied guanine oxidative lesion is 8-oxo-7,8-dihydroguanine (8-oxoG)^{11, 12} which is found at steady-state levels of 0.3-4.2 per 10^6 dG *in vivo*.¹³ 8-oxoG lesions form from the reaction of guanine with a variety of ROS and therefore serves as a biomarker to indicate the oxidative stress.¹⁴ Despite the minor structural difference compared to G (introduction of an *oxo* group to the C8 position and a hydrogen atom to the N7 position, atom highlighted in Figure 1.2), 8-oxoG has two different base pair patterns. 8-oxoG can base pair with the correct C or it can rotate about the glycosidic bond by 180° to mimic T functionally and form an 8-oxoG:A base pair,^{15, 16} with the latter leads to G→T mutation during the following round of replication (Figure 1.1). 8-oxoG has a high miscoding frequency *in vitro* but a low mutation frequency of $\sim 10\%$ *in vivo*, yielding less than 10% G→T mutations.^{17, 18} This

reduced mutation *in vivo* is explained by the extensive repair pathways dedicated to 8-oxoG, which will be described later in section 4.¹⁶



The reduction potential of 8-oxoG ($E_0 = 0.7$ V vs. NHE) is significantly lower than dG,^{19,}
²⁰ so it is labile to further oxidation that results in several secondary oxidation products or
 hydroperoxidized products (Figure 1.2).²¹ The initial step is a one-electron oxidation event converting
 8-oxoG to its radical cation, followed by subsequent addition of H₂O and deprotonation. A second
 one-electron oxidation and deprotonation leads to the product of 5-OH-8-oxoG. At pH > 7, 5-OH-
 8-oxoG will undergo a 1,2-shift and form the diastereomers of spiroiminodihydantoin (Sp);²² while
 at pH < 7, guanidinohydantoin (Gh) diastereomers will form as the hydrolysis and decarboxylation
 products.^{22, 23} While the Sp can be separated as individual diastereomers, Gh diastereomers exist
 as an equilibrium of the two species and cannot be isolated individually. While there has been
 speculation about the biological relevance of these hyperoxidized products, Sp was recently
 detected in genomic DNA from bacteria and mammalian cells. Hailer and co-workers detected the

Sp lesion in *E. coli* treated with chromate.²⁴ More recently, Mangerich and co-workers identified Sp in *Helicobacter hepaticus-infected* mice, which develop inflammation-mediated carcinogenesis at levels of 1-5 per 10⁸ nucleobases in genomic DNA.²⁵ Unlike the mild mutagenicity of 8-oxoG, Sp has showed 100% mutagenesis both *in vitro*^{26, 27} and *in vivo*,²⁸⁻³⁰ causing G→T and G→C mutations.



1.3 DNA Damage from Nucleotide Pool

When considering the possible ways to generate the oxidative DNA, it is very clear that ROS can directly act on genomic DNA, and is the focus of many studies in the field. However, it is also important to consider the contribution of oxidation of the building blocks used by DNA polymerase, the nucleotide pool. Early studies showed that the nucleotides are more susceptible to modification than the nucleobases located in the duplex form. Topal and Baker showed that when cells are treated with a methylating agent, the nucleotide pool is 190-13,000 fold more susceptible to modification than the DNA duplex.³¹ And more specifically for DNA oxidation, Kamiya and Kasai showed that dATP and dGTP from the nucleotide pool are 67-fold and 9-fold more easily oxidized respectively, compared with dA or dG in the duplex form.³² This modification preference is because that nucleobases in duplex form are less accessible for modification while nucleobases in nucleotides are lack of such protection. What makes the oxidation of the nucleotide pool more biologically significant is that cells encode an enzyme MutT/MTH1, a phosphatase that converts 8-oxodGTP to 8-oxodGMP,³³ and prevents the oxidized nucleotide from being the substrate for polymerase incorporation. The importance of removing 8-oxodGTP from the nucleotide pool is underscored by the fact that *E. coli* lacking MutT has a 100- to 10000-fold higher mutation rate compared with that of wild-type *E. coli*.³⁴ Indeed, it is known that 8-oxodGTP can be incorporated into DNA by several bacterial and mammalian polymerases.³⁵⁻³⁷ However, it remains unknown if polymerase can utilize the hyperoxidized nucleotides, such as dSpTP as a building block during DNA replication.

1.4 Repair of 8-oxoG and Sp Lesions

Although DNA is exposed to various oxidative damages to induce nucleobase modifications, the existence of DNA repair enzymes can alleviate the deleterious consequences. The MutT, MutM (also known as Fpg) and MutY enzymes in bacteria³⁸ or the corresponding human analogs³⁹ MTH1, OGG1 and MUTYH, are all responsible for repairing 8-oxoG. MTH1 hydrolyzes 8-oxodGTP to 8-oxodGMP and therefore prevents it from being incorporated into DNA.³³ OGG1 initiates the base excision repair (BER) pathway by excising the 8-oxoG lesions from DNA when they are paired with C and restore the canonical G:C base pair. Unlike OGG1 glycosylase, MUTYH can not remove 8-oxoG itself, but it selectively excises the A base in the opposite strand from an 8-oxoG:A base pair.⁴⁰ The relatively low mutation frequency of 8-oxoG in both *E. coli* and mammalian cells indicates the efficient repair of the lesion. For the hyperoxidized Sp lesion, unfortunately, MutT or other MutT-like enzymes cannot hydrolyze the dSpTP distereomers. It is not known now whether there is any other enzyme that can perform this specific function.⁴¹ *In vivo* studies have shown that Sp lesions are substrates for MutM, but the removal efficiency is higher when Sp is base-paired with C compared with A or G; however, the removal of Sp opposite of A is greater compared with the 8-oxoG opposite A.^{42, 43} While human OGG1 is responsible for the repair of 8-oxoG as MutM, it can not remove the Sp lesion, but the lesion is a substrate for yeast yOGG1, yOGG2 (Ntg1).⁴⁴ Some other glycosylases, such as endonuclease VIII (Nei) and its mammalian orthologues Nei-like (NEIL) family of enzymes (NEIL1, NEIL2 and NEIL3) can also remove the Sp lesion with different base pair preferences.^{43, 45-48} While most of the glycosylases discussed here showed activity towards the Sp lesion, they do not act on or have minimal activity on the 8-oxoG lesion. Finally, mismatch repair protein MutY cannot remove A from Sp:A base pairs.⁴² Therefore, the repair of lesions by various glycosylases

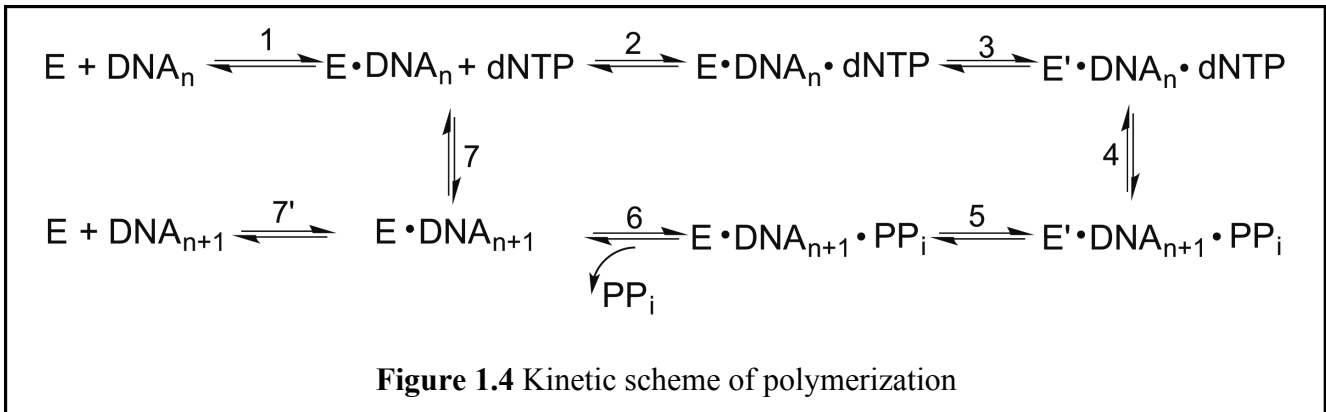
in BER depends on both the lesion type and the base pair pattern. Interestingly, despite the presence of the extensive repair pathways in cells, mutations are still observed when these lesions are formed in DNA,^{17, 29} suggesting incomplete repair of the lesions *in vivo*.

1.5 Chemical and Kinetic Mechanism of Polymerases

DNA replication involves multiple enzymes and processes in a cooperative and ordered manner. The polymerase is the catalytic core that is responsible for the chemistry of nucleotide addition. The polymerase incorporates four potential deoxynucleoside 5'-triphosphate (dNTP) onto a nascent strand via phosphoryl transfer in a 5' to 3' direction directed by the template with a distinct base pair pattern (A:T and G:C). The generally accepted catalytic mechanism involves two metal ions (Mg^{2+}) and several carboxylate residues from the active site of the polymerase. In detail, one Mg^{2+} activates the 3'-OH of the nascent strand to make it more nucleophilic, the electron rich oxygen on the 3' end then attacks the α -phosphate of the incoming nucleotide, creating a trigonal-bipyramidal pentacoordination transition state stabilized by the second Mg^{2+} by chelating the leaving oxygens on the β - and γ - phosphates,^{49, 50} followed by the leaving of pyrophosphate group.

In order to define the mechanism of how polymerase incorporates the correct nucleotide, it is important to understand the rates and equilibrium constants for all steps during the enzymatic catalytic cycle. For most DNA polymerases, the kinetic mechanism is described as follows (Figure 1.3).^{49, 51} The catalytic cycle starts with the polymerase binding to the DNA substrate to form an enzyme/DNA complex ($E \cdot DNA_n$) (step 1). Then an incoming nucleotide binds to the enzyme/DNA complex ($E \cdot DNA_n \cdot dNTP$) (step 2), where binding of the correct nucleotide will induce a conformational change ($E' \cdot DNA_n \cdot dNTP$) (step 3). It is hypothesized that this

conformational change will align the incoming nucleotide into a precise geometrical conformation allowing the following step to proceed. This step is also believed to be involved in the discrimination of incorrect dNTP incorporation when the geometry in the active site is disrupted. If the conformational change induces a close fit, chemistry step will occur, which is the phosphoryl transfer ($E' \cdot DNA_{n+1} \cdot PP_i$) (step 4). Following chemistry, there is another conformational change that relaxes the complex ($E \cdot DNA_{n+1} \cdot PP_i$) (step 5). Then pyrophosphate is released from the complex and one catalytic event is complete ($E \cdot DNA_{n+1} + PP_i$) (step 6). Depending on the processivity of the polymerase, it will remain bound to product DNA and translocate to the next position to continue the synthesis ($E \cdot DNA_{n+1} \cdot dNTP$) (step 7), or the polymerase will dissociate from the product DNA ($E + DNA_{n+1}$) (step 7') and continue the polymerization on another DNA substrate .

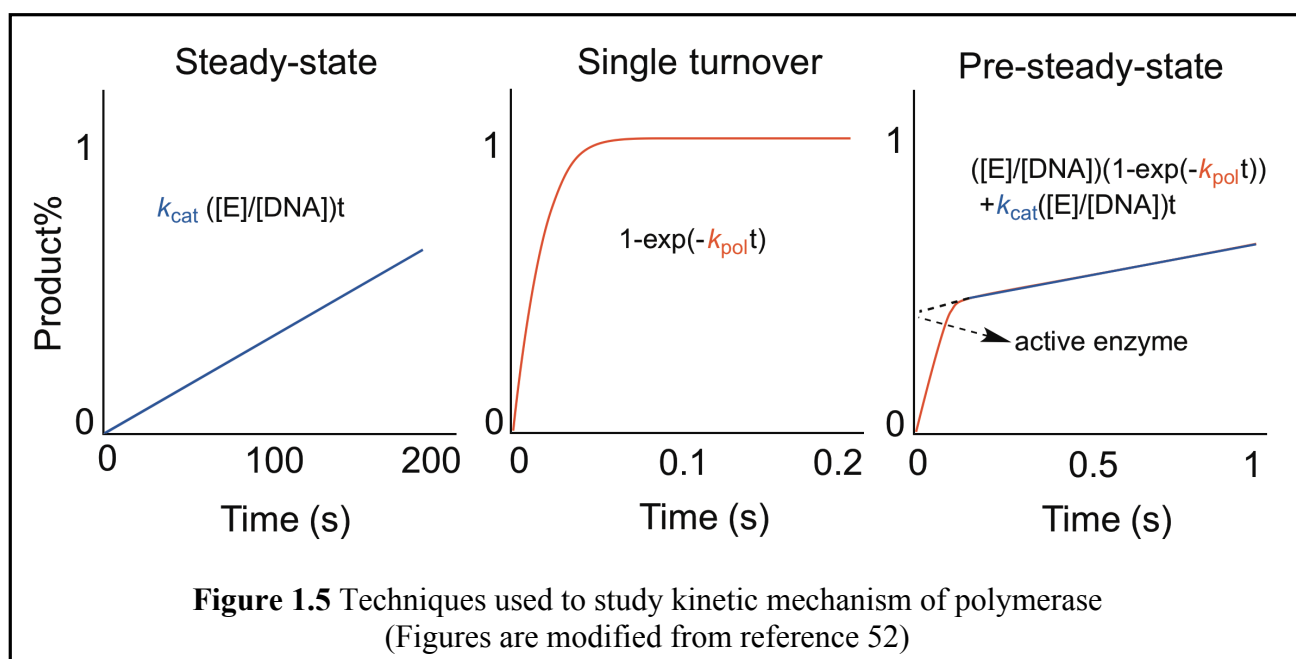


1.6 Techniques for Studying Nucleotide Incorporation Kinetics

Several kinetic methods have been developed to assess the polymerase activity of nucleotide incorporation. In particular, the use of a single nucleotide simplifies the product distribution but without considering the processivity of polymerase (Figure 1.4). Steady-state measurement is a well-studied technique for characterizing the nucleotide incorporation, which

use a large excess of DNA substrate compared to enzyme and the experiments are performed on a long time scale. Due to the large excess of DNA substrate compared to enzyme, many enzyme turnovers are needed to fully convert the DNA substrate to product. Therefore, the rate measured (k_{cat}) is limited by the slowest step of the reaction cycle, usually the product DNA released from polymerase. Also, the K_m measured from these kinetics is a complex mix of elementary rate constants that does not provide any information of nucleotide binding to the DNA-pol complex.⁵¹ Even though individual K_m and k_{cat} are usually uninformative on their own, the ratio of k_{cat}/K_m is used to measure the specificity of nucleotide selection. However, due to the limitation mentioned above, it fails to provide accurate parameters.⁵² A more useful method to study single nucleotide incorporation is the single turnover kinetics. The single turnover method becomes possible due to the ability of large-scale enzyme purification and instrumentation to perform very short timescale experiments. This technique requires large excess of enzyme compared to DNA, which is the opposite of the multiple turnover technique. In single turnover experiments, all the DNA substrate can bind to polymerase and convert to the DNA product in a single cycle, therefore eliminating the contribution from the slow DNA product release step. The observed rate, k_{pol} , reports the rate of the slowest step up to and including the chemistry step. Additional experiments are required to determine which of the pre-chemistry steps is rate-limiting. The observed rate of k_{pol} depends on the nucleotide concentration and the hyperbolic dependence of k_{pol} on dNTP concentration gives the K_d , which reflects the binding equilibrium that precedes the rate-limiting step. Evaluation of K_d and k_{pol} from single-turnover rapid quench methods has become a more prevailing technique for estimating the DNA polymerase fidelity as the ratio of k_{pol}/K_d .^{52, 53} A variation of single turnover kinetics is the pre steady state kinetics (or burst kinetics), where the concentration of polymerase is slightly lower than that of the DNA.⁵⁴ In this case, portions of the polymerase (or

active enzyme) bind to DNA and generate DNA product at the single-turnover rate. Conversion of the rest of the DNA needs the enzyme to dissociate from the product DNA product to recycle the reaction, therefore limiting the rate with the slowest product release step, just as in multiple turnover kinetics. If the single turnover and the multiple turnover rates are significantly different, a biphasic product formation curve will be observed, where the slope of each phase corresponds to k_{pol} and k_{cat} . Based on the amplitude of the fast phase, the active enzyme concentration can be calculated.



The above technique, combined with other methods, can be used to determine the rates along the steps during nucleotide incorporation, and therefore elucidate the kinetic mechanism of nucleotide incorporation by polymerases. Pre-steady state experiments show that *E. coli* DNA polymerase I has a burst phase (k_{pol}) 50 s^{-1} while the overall turnover rate k_{cat} is only $\sim 0.1 \text{ s}^{-1}$.⁵⁵ Some other polymerases also showed at least 50-fold slower rates when comparing k_{cat} to k_{pol} .^{56, 57} These results suggest that incorporation of a nucleotide is fast, followed by a rate-limiting product release step. In order to get more detailed information about the fast nucleotide incorporation step,

one must consider that the incorporation involves the nucleotide binding, a conformational change and phosphoryl transfer, early studies showed very little element effect when the oxygen on α -phosphate is replaced by sulfur, suggesting a step prior to the chemistry is fast and limits the rate of incorporation.⁵⁵ There is ongoing debate regarding whether the sulfur elemental effect can provide accurate information^{58, 59} and that the rate-limiting step might change according to the polymerase. For example, in a spectroscopic study of the nucleotide incorporation by polymerase beta (pol β), the authors showed that the rate of phosphoryl transfer is the rate-limiting step rather than any steps prior.⁶⁰ It is also possible that for a specific polymerase, the rate-limiting step will change when incorporating the correct versus the incorrect nucleotide.⁶¹

1.7 Introduction to Trinucleotide Repeat DNA

Regions of repetitive DNA, where a specific sequence motif repeats multiple times, constitute around 30% of the human genome.⁶²⁻⁶⁴ These regions are considered dynamic, often hotspots for nucleotide insertion or deletion, and more than 40 neurodegenerative or neuromuscular disorders have been linked to genomic instability of repetitive DNA.^{65, 66} Trinucleotide repeats (TNR), one example of repetitive DNA, have taken on special significance on the study in this field since the expansion of TNR underlies a large component of repeat-associated disorders (Table 1.1), including Huntington's disease (HD), myotonic dystrophy (DM1), fragile X syndrome type A (FRAXA), Friedreich's ataxia (FRDA), etc.⁶⁵⁻⁷²

Table 1.1 Representative diseases caused by TNR expansion

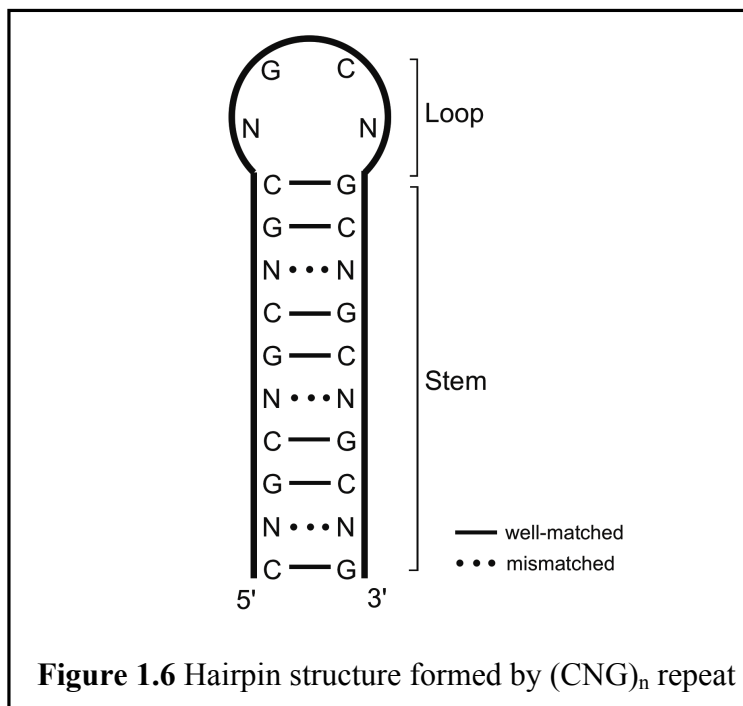
| Disorder | Gene Name (Protein) | Repeat | Putative Function | Normal Length | Disease Length |
|--------------------------------------|----------------------------|--------|----------------------------------|------------------|-------------------|
| Huntington's Disease (HD) | <i>HTT</i> (huntingtin) | CAG | Signaling Transport | 5-35 | >40 |
| Myotonic Dystrophy (DM1) | <i>DMPK</i> (DMPK) | CTG | RNA-mediated | 5-37 | >50 |
| Fragile X Syndrome Type A (FRAXA) | <i>FMRI</i> (FMRP) | CGG | Translational regulation | 6-60 | >200 |
| Friedreich's Ataxia (FRDA) | <i>FRDA</i> (frataxin) | GAA | Mitochondrial iron metabolism | 6-32 | >200 |

Despite cells having developed systems to prevent the TNR tract length from changing rapidly, once the TNR tract is longer than the crucial threshold length, an expanded TNR will be expected in the next generation. For example, HD is caused by the expansion of a (CAG)_n repeat tract in exon 1 of the *huntingtin* gene (*HTT*), located on the short arm of chromosome 4. Healthy individuals have 5-35 repeats, while a pre-mutation allele contains 36-39 repeats, which has an elevated chance to expand from one generation to another. Once the repeat tract is longer than 40 repeats, individuals will develop the symptoms of HD. Furthermore, HD is an autosomal dominant disorder, and therefore any offspring of the HD patients will have a 50% chance to be affected.⁷³ In addition, as the TNR becomes longer, the disease gets more severe and the age of onset is earlier for each successive generation, a phenomenon known as genetic anticipation.^{65, 71} The *HTT* encodes the polyglutamine (polyQ) containing huntingtin protein with a molecular weight of 348 kDa. Huntingtin protein is expressed ubiquitously in human and rodent, with the highest level in neurons of the central nervous system.⁷⁴ The relatively large size of the huntingtin protein has

prevented it from being crystalized and has complicated the mass spectrometry studies to elucidate the structure. As for now, there is still no clear understanding of the functions for the protein, but it is believed to be involved in embryonic development, clathrin-mediated endocytosis and neuronal gene transcriptional regulation, it also facilitates vesicular transport in axons, organizes the postsynaptic density and modulates the morphology of dendrites, last but not least, it is also an anti-apoptotic protein.⁷⁴⁻⁷⁶ In HD, expansion of the CAG repeat tract leads to the expression of the mutant huntingtin protein with expanded polyQ, which can mis-fold and aggregate. It is hypothesized that the pathology of HD is due to the loss of normal huntingtin protein functions and the incorrect interactions of the mutant huntingtin protein with other cellular components.⁷⁴⁻⁷⁶

1.8 TNR Expansion Mechanisms

While there is currently no effective treatment for the repeat expansion associated diseases, it is important to understand the molecular mechanism of the repeat expansion for the development of potential therapeutics. Several models have been proposed to explain the expansion mechanism and all of them lead to same conclusion that the TNR instability is the result of the formation of non-canonical secondary structure,⁷⁷⁻⁸⁰ such as hairpins, during DNA replication and repair process, as well as DNA transcription and DNA recombination.^{65, 68, 71} On the other hand, repeats that are not structure-prone are considerably more stable and do not show expansion. Single-stranded (CNG)_n TNR tracts are known to form hairpin structures (Figure 1.5),⁷⁹ which contain both well-matched Watson-Crick base pairs and mismatches. The stability of these (CNG)_n hairpins decrease in the order of CGG>CTG>CAG=CCG in physiological conditions and this difference is due to the energy contribution of the mismatch.⁶⁸



Many TNR loci are close to the replication origins, which suggests that TNR expansion might arise from the DNA replication.⁶⁶ Numerous *in vitro*⁸¹⁻⁸⁴ and *in vivo*^{85, 86} studies have demonstrated the propensity for TNR sequences to fold into hairpin structures. These hairpins impede replication by stalling of various DNA polymerase, which eventually lead to replication fork collapse (Figure 1.6).^{68, 87, 88} The stalling effect is also repeat length dependent and becomes more prominent when approaching the threshold repeat number of the disease state number.⁸⁷⁻⁹¹ Due to the fact that the formation of hairpins triggers the expansion of TNR, it requires DNA unwinding or even complete strand separation. This leads to the consideration of Okazaki fragment, a portion of the lagging strand template, which remains transiently single-stranded and it is therefore reasonable to suggest the TNR instability might come from the Okazaki region.⁶⁸ The replication-based expansion models also relies on the orientation and sequence of the DNA and this dependence is attributed to the stability of the proposed hairpin structure. For example, the study of the stability of $(CTG)_n/(CAG)_n$ repeats in bacterial plasmids showed that when the more stable or structure-prone CTG sequence were located in the lagging strand template, the repeat

DNA tract usually contracts; on the contrary, when CTG repeats were located in the nascent lagging strand, contraction was less frequent but expansion were observed.^{92, 93}

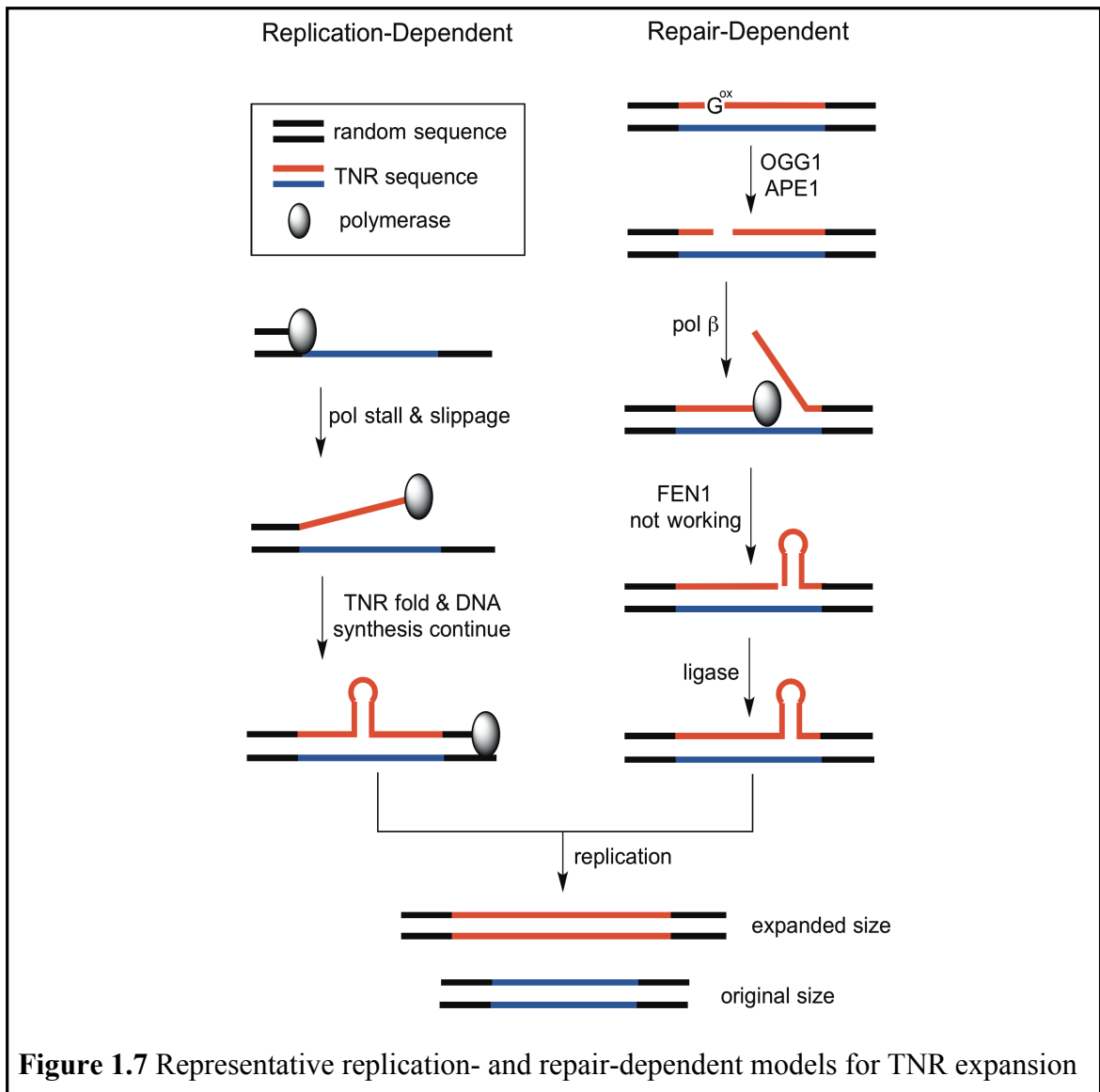


Figure 1.7 Representative replication- and repair-dependent models for TNR expansion

While the replication-based models can explain the TNR expansion between generations or in dividing cells, it is surprising that the largest expansions occur under non-dividing conditions, such as in the meiotically arrested oocytes or the terminally differentiated neurons. Evidence suggest that DNA repair machinery also contributes to the expansion of TNR, including DNA mismatch repair (MMR) and base excision repair (BER). In MMR, the MSH2-MSH3 dimer repairs

single-base insertions or small looped-out structures.⁹⁴ However, in a HD transgenic mouse model, where the MSH2 gene has been inactivated, the frequency of repeat expansion is significantly suppressed, suggesting MSH2 or MSH3 plays a role in TNR expansion.^{95, 96} One possible explanation involves the structural properties of hairpins formed by CNG repeats. Mismatch repair proteins MSH2-MSH3 can recognize and bind to the N•N mismatches, and it has been shown that the binding affinity is stronger for longer CNG repeats containing more N•N mismatches.⁹⁷ The strong binding of the repair proteins therefore stabilizes the hairpin structure and also significantly reduces the ability of MSH2-MSH3 to hydrolyze ATP,^{72, 97} an essential function involved in canonical MMS repair pathway. The incorporation of the hairpin is completed by a proposed a nick filling mechanism to produce an expanded repeat tract.⁷¹

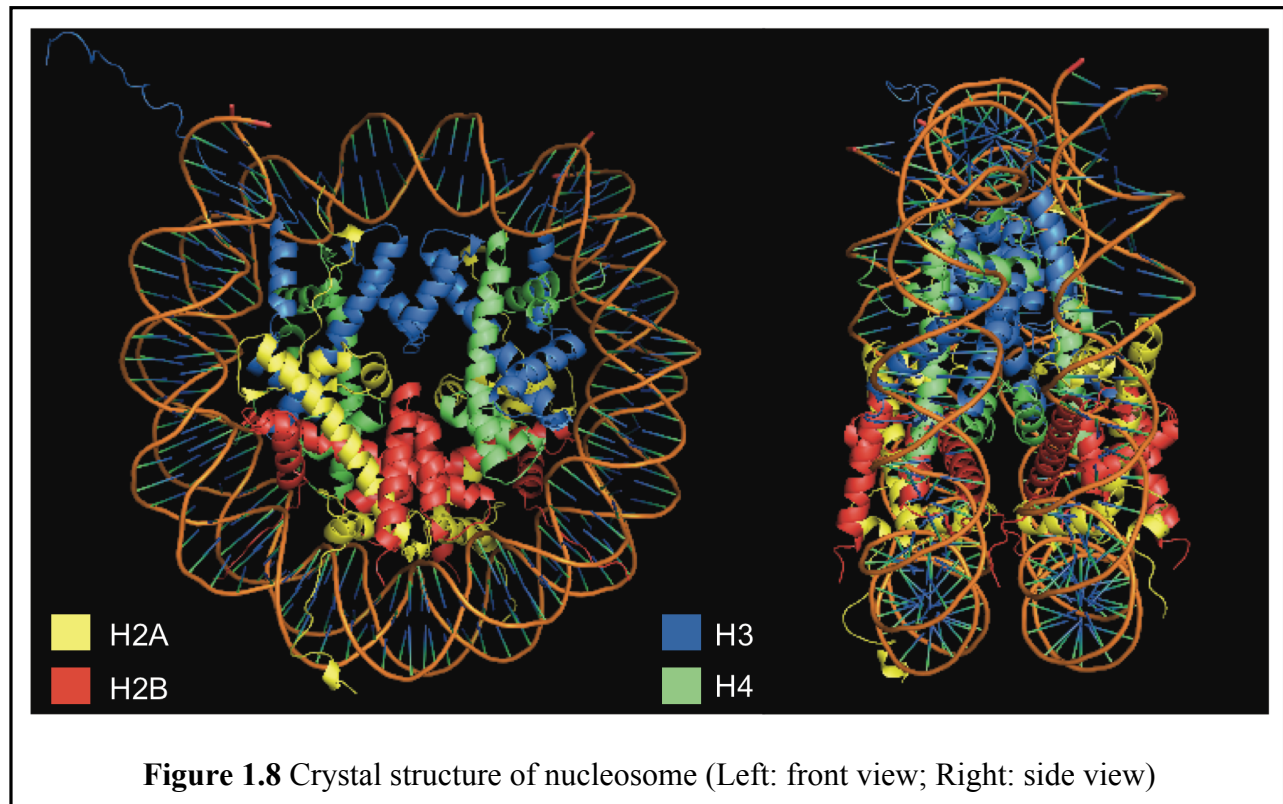
TNR expansion has also been shown to depend on oxidative damage. A HD mouse model showed an age-dependent TNR expansion in somatic cells with accumulated level of 8-oxoG.⁹⁸ When the function of OGG1, a glycosylase initiates the BER pathway, was inactivated in the HD mouse model, TNR expansion decreased. These results suggest that repair of oxidative damage in the CAG repeats will also lead to the TNR expansion. In this model (Figure 1.6), OGG1 removes the 8-oxoG base and APE1 creates a nick on the 5' side. Pol β performs multi-nucleotide synthesis and generates a (CAG)_n flap, which can fold into a hairpin structure. The formation of hairpin structure is refractory to cleavage by the downstream endonuclease, FEN1. As a result of FEN1's inability to cleave the 5'-hairpin, extra TNR repeats is ligated into the sequence. Our lab has also shown that the guanine in the hairpin loop is more accessible for further oxidation and this will lead to an iterative long-patch BER induced TNR expansion.⁹⁹

Other models have also been proposed to explain the repeat instability and TNR expansion, including DNA recombination, nucleotide excision repair, and RNA transcription, and have been

reviewed elsewhere.^{65, 66, 71, 100}

1.9 Nucleosome -- Basic Unit for DNA Packaging

Despite the complexity of eukaryotic chromatin, the high order structure is organized by repeating the same unit called a nucleosome. A nucleosome consists of the nucleosome core particle (NCP) and the linker region (linker DNA and linker histone). The NCP contains two copies of each histone protein, H2A, H2B, H3 and H4. These histone proteins assemble into two copies of an H2A-H2B dimer and one copy of an (H3-H4)₂ tetramer. From the crystal structure of nucleosome, one can see that the two H2A-H2B dimers are anchored to the (H3-H4)₂ tetramer, forming a histone octamer. The octamer is then wrapped around by 145-147 base pairs (bp) DNA in about 1.6 turns of a left-handed superhelix to form the NCP (Figure 1.7).^{101, 102}

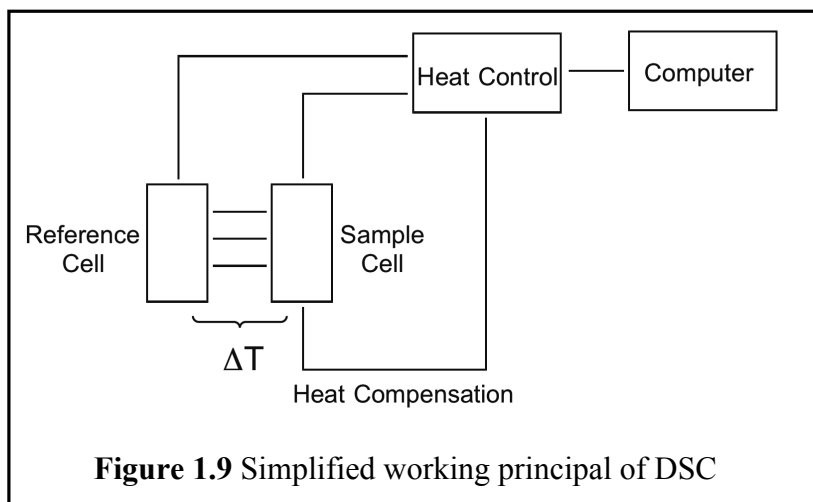


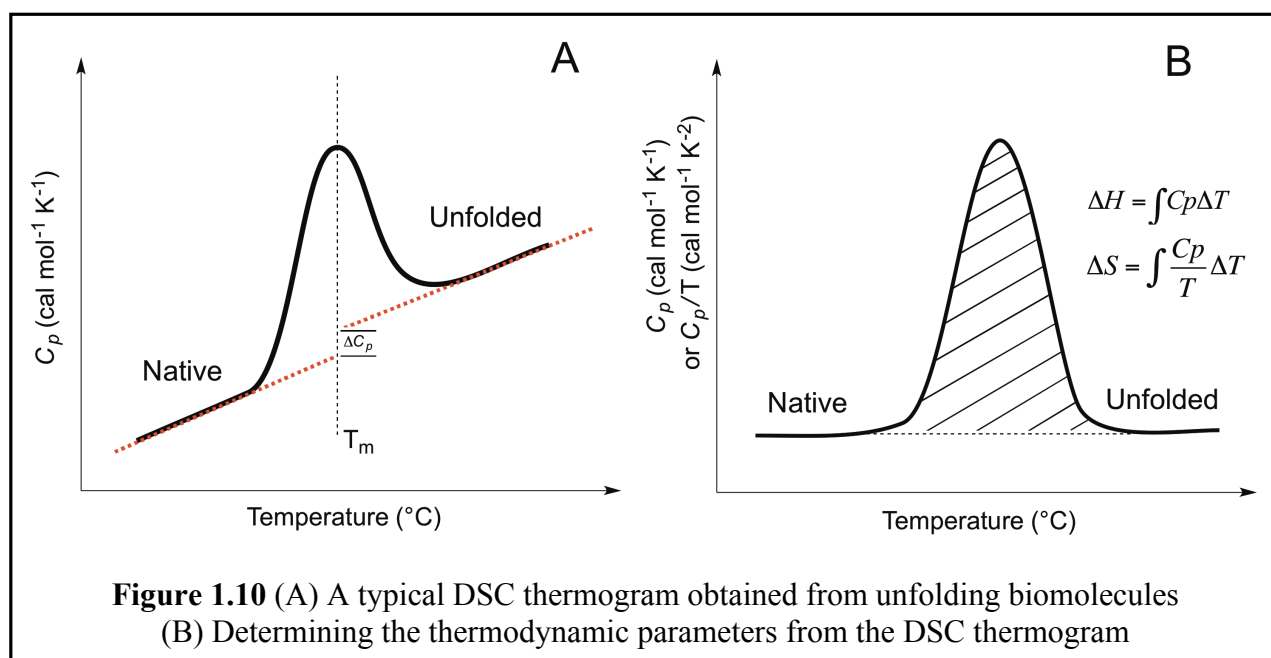
NCPs are then connected and stabilized by linker DNA and a fifth linker histone H1.¹⁰¹ This organization allows the genomic DNA to fit into the nucleus and allows for the regulation of various cellular processes, such as DNA replication, transcription and repair. Histone proteins are highly alkaline proteins and will undergo various post-translational modifications, including methylation, acetylation, phosphorylation, ubiquitination, SUMOylation and others.¹⁰³ These modifications can occur on both histone tail region (unstructured region of histone proteins that usually point away from the center of the histone octamer) or in the histone core (structured region of histone proteins). These distinct histone modifications, on one or multiple histones, act sequentially or in combination to form a “histone code” that is read by other proteins to bring about distinct downstream events, expanding the idea of the genetic code.¹⁰⁴ One example is that the acetylation of histone proteins¹⁰⁵ will neutralize the charge on the lysine residue, therefore reducing the interactions between the positively charged histones and the negatively charged DNA backbone. By this mechanism, the condensed chromatin is converted into a more relaxed form for gene transcription. This relaxation can be reversed by histone deacetylase (HDAC), which removes the acetyl group from histones.

1.10 Introduction to Differential Scanning Calorimetry

The characterization of thermodynamic parameters of unfolding of biomolecules using direct calorimetric methods dates back as early as the 1960s'. Technology advances has made it commercially available as a routine technique for biophysical characterization.¹⁰⁶ Calorimetric studies have been widely used in the characterization the lipids, proteins, nucleic acids, as well as drug discovery development.^{107, 108} Differential scanning calorimetry (DSC) is a thermoanalytical technique to analyze samples, typically on a milligram scale. A DSC instrument contains two cells,

a reference cell and a sample cell, and a conventional forward-scanning (heating) mode that monitors a small temperature difference between the reference cell that is filled with buffer and the sample cell filled with biomolecules of interest in an identical buffer. During the forward-scanning process at constant rate (typically 1 °C/min or less), the temperature increases and thermally induced transition occurs in the sample cell, resulting in heat being absorbed. This produces a temperature difference (ΔT) relative to the reference cell, and this difference is compensated for by the power source to its initial value and recorded by a computer. This compensated heat is proportional to the excess heat capacity of the thermally induced transition (Figure 1.8). Usually a buffer/buffer scan is performed prior to the buffer/sample scan, and serves as a background signal to account for the instrumental difference between the two cells. The background signal determined can then be subtracted from the buffer/sample scan. Additionally, the native and unfolded biomolecules usually exhibit different heat capacities, therefore creating two baselines (either pre- to or post-transition) to account for it (Figure 1.9A). The difference of heat capacity (ΔC_p) at the transition is important to consider when analyzing the transition. After baseline correction and concentration normalization, the area under the transition can be integrated to give a direct measurement of enthalpy change (ΔH) and melting temperature (T_m) (Figure 1.9B).





A simplified model of equilibrium unfolding of biomolecules assumes that the molecule only belongs in two thermodynamic states: the folded state (or native state) and the unfolded state. This “all-or-none” model was first proposed by Tim Anson in 1945¹⁰⁹ and is believed to hold true for the unfolding of small molecules.^{110, 111} In reality, larger molecules may exhibit multiple intermediate states during the unfolding process. A spectrometric method can be used to obtain the enthalpy derived from equilibrium constants by means of the van’t Hoff equation under the assumption of the two-state model; on the other hand, calorimetry is a direct measurement of heat change associated with the unfolding process and therefore it is model-independent. As the results, it is possible to determine whether an unfolding process follows the two-state model based on the enthalpy values calculated from the spectroscopic (ΔH_{VH}) and calorimetric (ΔH_{Cal}) experiments. In detail, when $\Delta H_{VH}/\Delta H_{Cal} = 1$, two-state model holds true during the unfolding process; when $\Delta H_{VH}/\Delta H_{Cal} < 1$, two-state model is no longer followed and intermediates may populate during the unfolding process, which the spectrometric method can not capture. In this case, it is an advantage

to utilize the DSC to acquire thermodynamic parameters associated with the unfolding process of biomolecules.

1.11 Concluding Remarks and Gap in Knowledge

When considering the types of DNA mutations, the modification of nucleobases may change their canonical base pair pattern and therefore introduce point mutations. Sp lesions have been widely investigated for its mutation profile both *in vitro* and *in vivo*, and the results showed that Sp is highly mutagenic with a mutation frequency of nearly 100%. While most of the studies of Sp focus on the oxidation of nucleobases within a DNA sequence, it is still not clear whether the generation of Sp from the nucleotide pool will contribute to the mutagenicity of Sp. Therefore, it is important to know how polymerase utilizes the Sp from the nucleotide pool as a substrate during DNA replication. This involves two aspects: one of which is whether there will be any difference in base pair pattern when Sp lesion locates in DNA or as a nucleotide; the other is how efficiently the polymerase can incorporate the Sp nucleotide. These questions need to be answered in order to give a thorough understanding of the mutagenicity of Sp. The first part of this thesis gives a detailed discussion about the Sp nucleotide as a substrate for polymerase incorporation compared with its precursors dGTP and 8-oxodGTP.

From a different perspective of DNA mutation, the change of DNA sequence may not be due to a single base, such as during the TNR expansion, discussed above. Although a non-canonical base pair pattern is not involved, the length of the DNA is changed and sometimes is changed significantly. Formation of hairpin structures are involved in various proposed models for the expansion mechanism and the thermodynamic stability of these hairpins are thought to be important to promote the expansion. In the second part of this thesis, we elucidate how (CTG)_n

and (CAG)_n repeat number affect the thermodynamic stability of the hairpins by differential scanning calorimetry (DSC). In addition, the kinetic stability of complementary (CTG)_n and (CAG)_n hairpins will also be evaluated when changing of the repeat number.

As a final consideration, one must consider the characterization of (CAG)_n/(CTG)_n containing DNA in a more biologically relevant environment. We examined the thermodynamic behaviors of DNA with various number of (CAG)_n/(CTG)_n repeat in the nucleosomal length, we also reconstituted these DNA into nucleosomes to see how these expanded TNR affect the stability of nucleosomes by DSC.

1.12 Reference

- (1) Watson, J. D., and Crick, F. H. C. (1953) Molecular Structure of Nucleic Acids: A Structure for Deoxyribose Nucleic Acid. *Nature* 171, 737-738.
- (2) Crick, F. H. C., Wang, J. C., and Bauer, W. R. (1979) Is DNA really a double helix? *J Mol Biol* 129, 449-461.
- (3) Gates, K. S. (2009) An overview of chemical processes that damage cellular DNA: spontaneous hydrolysis, alkylation, and reactions with radicals. *Chem Res Toxicol* 22, 1747-1760.
- (4) Hussain, S. P., Hofseth, L. J., and Harris, C. C. (2003) Radical causes of cancer. *Nat Rev Cancer* 3, 276-285.
- (5) Niles, J. C., Wishnok, J. S., and Tannenbaum, S. R. (2006) Peroxynitrite-induced oxidation and nitration products of guanine and 8-oxoguanine: structures and mechanisms of product formation. *Nitric Oxide* 14, 109-121.

- (6) Regulus, P., Duroux, B., Bayle, P. A., Favier, A., Cadet, J., and Ravanat, J. L. (2007) Oxidation of the sugar moiety of DNA by ionizing radiation or bleomycin could induce the formation of a cluster DNA lesion. *Proc Natl Acad Sci U S A* 104, 14032-14037.
- (7) Regulus, P., Spessotto, S., Gateau, M., Cadet, J., Favier, A., and Ravanat, J. L. (2004) Detection of new radiation-induced DNA lesions by liquid chromatography coupled to tandem mass spectrometry. *Rapid Commun Mass Spectrom* 18, 2223-2228.
- (8) Candeias, L. P., O'Neill, P., Jones, G. D. D., and Steenken, S. (1992) Ionization of Polynucleotides and DNA in Aqueous Solution by 193nm Pulsed Laser Light: Identification of Base-derived Radicals. *International Journal of Radiation Biology* 61, 15-20.
- (9) Pouget, J. P., Douki, T., Richard, M. J., and Cadet, J. (2000) DNA Damage Induced in Cells by γ and UVA Radiation As Measured by HPLC/GC-MS and HPLC-EC and Comet Assay. *Chemical Research in Toxicology* 13, 541-549.
- (10) Steenken, S., and Jovanovic, S. V. (1997) How Easily Oxidizable Is DNA? One-Electron Reduction Potentials of Adenosine and Guanosine Radicals in Aqueous Solution. *J Am Chem Soc* 119, 617-618.
- (11) Neeley, W. L., and Essigmann, J. M. (2006) Mechanisms of formation, genotoxicity, and mutation of guanine oxidation products. *Chem Res Toxicol* 19, 491-505.
- (12) Burrows, C. J., and Muller, J. G. (1998) Oxidative Nucleobase Modifications Leading to Strand Scission. *Chem Rev* 98, 1109-1152.
- (13) Moller, P., Cooke, M. S., Collins, A., Olinski, R., Rozalski, R., and Loft, S. (2012) Harmonising measurements of 8-oxo-7,8-dihydro-2'-deoxyguanosine in cellular DNA and urine. *Free Radic Res* 46, 541-553.

- (14) Klaunig, J. E., and Kamendulis, L. M. (2004) The role of oxidative stress in carcinogenesis. *Annu Rev Pharmacol Toxicol* 44, 239-267.
- (15) Batra, V. K., Shock, D. D., Beard, W. A., McKenna, C. E., and Wilson, S. H. (2012) Binary complex crystal structure of DNA polymerase beta reveals multiple conformations of the templating 8-oxoguanine lesion. *Proc Natl Acad Sci U S A* 109, 113-118.
- (16) David, S. S., O'Shea, V. L., and Kundu, S. (2007) Base-excision repair of oxidative DNA damage. *Nature* 447, 941-950.
- (17) Cheng, K. C., Cahill, D. S., Kasai, H., Nishimura, S., and Loeb, L. A. (1992) 8-Hydroxyguanine, an abundant form of oxidative DNA damage, causes G----T and A----C substitutions. *J Biol Chem* 267, 166-172.
- (18) Klein, J. C., Bleeker, M. J., Saris, C. P., Roelen, H. C., Brugghe, H. F., van den Elst, H., van der Marel, G. A., van Boom, J. H., Westra, J. G., Kriek, E., and et al. (1992) Repair and replication of plasmids with site-specific 8-oxodG and 8-AAFdG residues in normal and repair-deficient human cells. *Nucleic Acids Res* 20, 4437-4443.
- (19) Berger, M., Anselmino, C., Mouret, J.-F., and Cadet, J. (1990) High Performance Liquid Chromatography-Electrochemical Assay for Monitoring the Formation of 8-Oxo-7,8-dihydroadenine and its Related 2' -Deoxyribonucleoside. *Journal of Liquid Chromatography* 13, 929-940.
- (20) Yanagawa, H., Ogawa, Y., and Ueno, M. (1992) Redox ribonucleosides. Isolation and characterization of 5-hydroxyuridine, 8-hydroxyguanosine, and 8-hydroxyadenosine from *Torula* yeast RNA. *J Biol Chem* 267, 13320-13326.

- (21) Steenken, S., Jovanovic, S. V., Bietti, M., and Bernhard, K. (2000) The trap depth (in DNA) of 8-oxo-7,8-dihydro-2' deoxyguanosine as derived from electron-transfer equilibria in aqueous solution. *J Am Chem Soc* 122, 2373-2374.
- (22) Alshykhly, O. R., Fleming, A. M., and Burrows, C. J. (2015) Guanine oxidation product 5-carboxamido-5-formamido-2-iminohydantoin induces mutations when bypassed by DNA polymerases and is a substrate for base excision repair. *Chem Res Toxicol* 28, 1861-1871.
- (23) Ye, Y., Muller, J. G., Luo, W., Mayne, C. L., Shallop, A. J., Jones, R. A., and Burrows, C. J. (2003) Formation of ¹³C-, ¹⁵N-, and ¹⁸O-labeled guanidinohydantoin from guanosine oxidation with singlet oxygen. Implications for structure and mechanism. *J Am Chem Soc* 125, 13926-13927.
- (24) Hailer, M. K., Slade, P. G., Martin, B. D., and Sugden, K. D. (2005) Nei deficient Escherichia coli are sensitive to chromate and accumulate the oxidized guanine lesion spiroiminodihydantoin. *Chem Res Toxicol* 18, 1378-1383.
- (25) Mangerich, A., Knutson, C. G., Parry, N. M., Muthupalani, S., Ye, W., Prestwich, E., Cui, L., McFaline, J. L., Mobley, M., Ge, Z., Taghizadeh, K., Wishnok, J. S., Wogan, G. N., Fox, J. G., Tannenbaum, S. R., and Dedon, P. C. (2012) Infection-induced colitis in mice causes dynamic and tissue-specific changes in stress response and DNA damage leading to colon cancer. *Proc Natl Acad Sci U S A* 109, E1820-1829.
- (26) Duarte, V., Muller, J. G., and Burrows, C. J. (1999) Insertion of dGMP and dAMP during in vitro DNA synthesis opposite an oxidized form of 7,8-dihydro-8-oxoguanine. *Nucleic Acids Res* 27, 496-502.

- (27) Korniyushyna, O., Berges, A. M., Muller, J. G., and Burrows, C. J. (2002) In Vitro Nucleotide Misinsertion Opposite the Oxidized Guanosine Lesions Spiroiminodihydantoin and Guanidinohydantoin and DNA Synthesis Past the Lesions Using *Escherichia coli* DNA Polymerase I (Klenow Fragment)†. *Biochemistry-Us* 41, 15304-15314.
- (28) Henderson, P. T., Delaney, J. C., Muller, J. G., Neeley, W. L., Tannenbaum, S. R., Burrows, C. J., and Essigmann, J. M. (2003) The hydantoin lesions formed from oxidation of 7,8-dihydro-8-oxoguanine are potent sources of replication errors in vivo. *Biochemistry-Us* 42, 9257-9262.
- (29) Delaney, S., Delaney, J. C., and Essigmann, J. M. (2007) Chemical-biological fingerprinting: probing the properties of DNA lesions formed by peroxyxynitrite. *Chem Res Toxicol* 20, 1718-1729.
- (30) Delaney, S., Neeley, W. L., Delaney, J. C., and Essigmann, J. M. (2007) The substrate specificity of MutY for hyperoxidized guanine lesions in vivo. *Biochemistry-Us* 46, 1448-1455.
- (31) Topal, M. D., and Baker, M. S. (1982) DNA precursor pool: a significant target for N-methyl-N-nitrosourea in C3H/10T1/2 clone 8 cells. *Proc Natl Acad Sci U S A* 79, 2211-2215.
- (32) Kamiya, H., and Kasai, H. (1995) Formation of 2-hydroxydeoxyadenosine triphosphate, an oxidatively damaged nucleotide, and its incorporation by DNA polymerases. Steady-state kinetics of the incorporation. *J Biol Chem* 270, 19446-19450.
- (33) Maki, H., and Sekiguchi, M. (1992) MutT protein specifically hydrolyses a potent mutagenic substrate for DNA synthesis. *Nature* 355, 273-275.

- (34) Yanofsky, C., Cox, E. C., and Horn, V. (1966) The unusual mutagenic specificity of an E. Coli mutator gene. *Proc Natl Acad Sci U S A* 55, 274-281.
- (35) Pavlov, Y. I., Minnick, D. T., Izuta, S., and Kunkel, T. A. (1994) DNA Replication Fidelity with 8-Oxodeoxyguanosine Triphosphate. *Biochemistry-U S* 33, 4695-4701.
- (36) Einolf, H. J., and Guengerich, F. P. (2001) Fidelity of nucleotide insertion at 8-oxo-7,8-dihydroguanine by mammalian DNA polymerase delta. Steady-state and pre-steady-state kinetic analysis. *J Biol Chem* 276, 3764-3771.
- (37) Hanes, J. W., Thal, D. M., and Johnson, K. A. (2006) Incorporation and replication of 8-oxo-deoxyguanosine by the human mitochondrial DNA polymerase. *J Biol Chem* 281, 36241-36248.
- (38) Michaels, M. L., and Miller, J. H. (1992) The GO system protects organisms from the mutagenic effect of the spontaneous lesion 8-hydroxyguanine (7,8-dihydro-8-oxoguanine). *J Bacteriol* 174, 6321-6325.
- (39) Barnes, D. E., and Lindahl, T. (2004) Repair and genetic consequences of endogenous DNA base damage in mammalian cells. *Annu Rev Genet* 38, 445-476.
- (40) Manuel, R. C., Hitomi, K., Arvai, A. S., House, P. G., Kurtz, A. J., Dodson, M. L., McCullough, A. K., Tainer, J. A., and Lloyd, R. S. (2004) Reaction intermediates in the catalytic mechanism of Escherichia coli MutY DNA glycosylase. *J Biol Chem* 279, 46930-46939.
- (41) Hori, M., Suzuki, T., Minakawa, N., Matsuda, A., Harashima, H., and Kamiya, H. (2011) Mutagenicity of secondary oxidation products of 8-oxo-7,8-dihydro-2'-deoxyguanosine 5'-triphosphate (8-hydroxy-2'- deoxyguanosine 5'-triphosphate). *Mutat Res* 714, 11-16.

- (42) Leipold, M. D., Muller, J. G., Burrows, C. J., and David, S. S. (2000) Removal of hydantoin products of 8-oxoguanine oxidation by the *Escherichia coli* DNA repair enzyme, FPG. *Biochemistry-Us* 39, 14984-14992.
- (43) Hazra, T. K., Muller, J. G., Manuel, R. C., Burrows, C. J., Lloyd, R. S., and Mitra, S. (2001) Repair of hydantoins, one electron oxidation product of 8-oxoguanine, by DNA glycosylases of *Escherichia coli*. *Nucleic Acids Res* 29, 1967-1974.
- (44) Leipold, M. D., Workman, H., Muller, J. G., Burrows, C. J., and David, S. S. (2003) Recognition and removal of oxidized guanines in duplex DNA by the base excision repair enzymes hOGG1, yOGG1, and yOGG2. *Biochemistry-Us* 42, 11373-11381.
- (45) Bandaru, V. (2002) A novel human DNA glycosylase that removes oxidative DNA damage and is homologous to *Escherichia coli* endonuclease VIII. *DNA Repair* 1, 517-529.
- (46) Hazra, T. K., Izumi, T., Boldogh, I., Imhoff, B., Kow, Y. W., Jaruga, P., Dizdaroglu, M., and Mitra, S. (2002) Identification and characterization of a human DNA glycosylase for repair of modified bases in oxidatively damaged DNA. *Proc Natl Acad Sci U S A* 99, 3523-3528.
- (47) Hazra, T. K., Kow, Y. W., Hatahet, Z., Imhoff, B., Boldogh, I., Mokkalapati, S. K., Mitra, S., and Izumi, T. (2002) Identification and characterization of a novel human DNA glycosylase for repair of cytosine-derived lesions. *J Biol Chem* 277, 30417-30420.
- (48) Morland, I. (2002) Human DNA glycosylases of the bacterial Fpg/MutM superfamily: an alternative pathway for the repair of 8-oxoguanine and other oxidation products in DNA. *Nucleic Acids Res* 30, 4926-4936.
- (49) Berdis, A. J. (2009) Mechanisms of DNA polymerases. *Chem Rev* 109, 2862-2879.

- (50) Steitz, T. A. (1999) DNA Polymerases: Structural Diversity and Common Mechanisms. *J Biol Chem* 274, 17395-17398.
- (51) Joyce, C. M. (2010) Techniques used to study the DNA polymerase reaction pathway. *Biochim Biophys Acta* 1804, 1032-1040.
- (52) Tsai, Y. C., and Johnson, K. A. (2006) A new paradigm for DNA polymerase specificity. *Biochemistry-Us* 45, 9675-9687.
- (53) Lee, H. R., Helquist, S. A., Kool, E. T., and Johnson, K. A. (2008) Importance of hydrogen bonding for efficiency and specificity of the human mitochondrial DNA polymerase. *J Biol Chem* 283, 14402-14410.
- (54) Johnson, K. A. (1995) Rapid quench kinetic analysis of polymerases, adenosinetriphosphatases, and enzyme intermediates. *Methods Enzymol.* 249, 38-61.
- (55) Kuchta, R. D., Benkovic, P., and Benkovic, S. J. (1988) Kinetic mechanism whereby DNA polymerase I (Klenow) replicates DNA with high fidelity. *Biochemistry-Us* 27, 6716-6725.
- (56) Wong, I., Patel, S. S., and Johnson, K. A. (1991) An induced-fit kinetic mechanism for DNA replication fidelity: direct measurement by single-turnover kinetics. *Biochemistry-Us* 30, 526-537.
- (57) Capson, T. L., Peliska, J. A., Kaboord, B. F., Frey, M. W., Lively, C., Dahlberg, M., and Benkovic, S. J. (1992) Kinetic characterization of the polymerase and exonuclease activities of the gene 43 protein of bacteriophage T4. *Biochemistry-Us* 31, 10984-10994.
- (58) Herschlag, D., Piccirilli, J. A., and Cech, T. R. (1991) Ribozyme-catalyzed and nonenzymic reactions of phosphate diesters: rate effects upon substitution of sulfur for a nonbridging phosphoryl oxygen atom. *Biochemistry-Us* 30, 4844-4854.

- (59) Showalter, A. K., and Tsai, M.-D. (2002) A Reexamination of the Nucleotide Incorporation Fidelity of DNA Polymerases[†]. *Biochemistry-U.S.* 41, 10571-10576.
- (60) Arndt, J. W., Gong, W., Zhong, X., Showalter, A. K., Liu, J., Dunlap, C. A., Lin, Z., Paxson, C., Tsai, M.-D., and Chan, M. K. (2001) Insight into the Catalytic Mechanism of DNA Polymerase β : Structures of Intermediate Complexes^{†,‡}. *Biochemistry-U.S.* 40, 5368-5375.
- (61) Joyce, C. M., and Benkovic, S. J. (2004) DNA polymerase fidelity: kinetics, structure, and checkpoints. *Biochemistry-U.S.* 43, 14317-14324.
- (62) Beckman, J. S., and Weber, J. L. (1992) Survey of human and rat microsatellites. *Genomics* 12, 627-631.
- (63) Jurka, J., and Pethiyagoda, C. (1995) Simple repetitive DNA sequences from primates: Compilation and analysis. *Journal of Molecular Evolution* 40, 120-126.
- (64) Toth, G. (2000) Microsatellites in Different Eukaryotic Genomes: Survey and Analysis. *Genome Research* 10, 967-981.
- (65) Pearson, C. E., Nichol Edamura, K., and Cleary, J. D. (2005) Repeat instability: mechanisms of dynamic mutations. *Nat Rev Genet* 6, 729-742.
- (66) Lopez Castel, A., Cleary, J. D., and Pearson, C. E. (2010) Repeat instability as the basis for human diseases and as a potential target for therapy. *Nat Rev Mol Cell Biol* 11, 165-170.
- (67) Bowater, R. P., and Wells, R. D. (2001) The intrinsically unstable life of DNA triplet repeats associated with human hereditary disorders. *Prog Nucleic Acid Res Mol Biol* 66, 159-202.
- (68) Mirkin, S. M. (2007) Expandable DNA repeats and human disease. *Nature* 447, 932-940.

- (69) Orr, H. T., and Zoghbi, H. Y. (2007) Trinucleotide repeat disorders. *Annu Rev Neurosci* 30, 575-621.
- (70) Kovtun, I. V., and McMurray, C. T. (2008) Features of trinucleotide repeat instability in vivo. *Cell Res* 18, 198-213.
- (71) McMurray, C. T. (2010) Mechanisms of trinucleotide repeat instability during human development. *Nat Rev Genet* 11, 886-886.
- (72) Iyer, R. R., Pluciennik, A., Napierala, M., and Wells, R. D. (2015) DNA triplet repeat expansion and mismatch repair. *Annu Rev Biochem* 84, 199-226.
- (73) Group, H. s. D. C. R. (1993) A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington's disease chromosomes. *Cell* 72, 971-983.
- (74) Cattaneo, E., Zuccato, C., and Tartari, M. (2005) Normal huntingtin function: an alternative approach to Huntington's disease. *Nat Rev Neurosci* 6, 919-930.
- (75) Zuccato, C., Valenza, M., and Cattaneo, E. (2010) Molecular mechanisms and potential therapeutical targets in Huntington's disease. *Physiol Rev* 90, 905-981.
- (76) Harjes, P., and Wanker, E. E. (2003) The hunt for huntingtin function: interaction partners tell many different stories. *Trends Biochem Sci* 28, 425-433.
- (77) Wells, R. D. (2007) Non-B DNA conformations, mutagenesis and disease. *Trends Biochem Sci* 32, 271-278.
- (78) Mirkin, S. M. (2006) DNA structures, repeat expansions and human hereditary disorders. *Curr Opin Struct Biol* 16, 351-358.
- (79) Kiliszek, A., and Rypniewski, W. (2014) Structural studies of CNG repeats. *Nucleic Acids Res* 42, 8189-8199.

- (80) McMurray, C. T. (1999) DNA secondary structure: A common and causative factor for expansion in human disease. *P Natl Acad Sci USA* 96, 1823-1825.
- (81) Amrane, S., Sacca, B., Mills, M., Chauhan, M., Klump, H. H., and Mergny, J. L. (2005) Length-dependent energetics of (CTG)(n) and (CAG)(n) trinucleotide repeats. *Nucleic Acids Res* 33, 4065-4077.
- (82) Amrane, S., and Mergny, J. L. (2006) Length and pH-dependent energetics of (CCG)(n) and (CGG)(n) trinucleotide repeats. *Biochimie* 88, 1125-1134.
- (83) Figueroa, A. A., Cattie, D., and Delaney, S. (2011) Structure of Even/Odd Trinucleotide Repeat Sequences Modulates Persistence of Non-B Conformations and Conversion to Duplex. *Biochemistry-Us* 50, 4441-4450.
- (84) Ohshima, K., and Wells, R. D. (1997) Hairpin formation during DNA synthesis primer realignment in vitro in triplet repeat sequences from human hereditary disease genes. *J Biol Chem* 272, 16798-16806.
- (85) Gacy, A. M., Goellner, G., Juranic, N., Macura, S., and McMurray, C. T. (1995) Trinucleotide Repeats That Expand in Human-Disease Form Hairpin Structures in-Vitro. *Cell* 81, 533-540.
- (86) Moore, H., Greenwell, P. W., Liu, C. P., Arnheim, N., and Petes, T. D. (1999) Triplet repeats form secondary structures that escape DNA repair in yeast. *P Natl Acad Sci USA* 96, 1504-1509.
- (87) Samadashwily, G. M., Raca, G., and Mirkin, S. M. (1997) Trinucleotide repeats affect DNA replication in vivo. *Nat Genet* 17, 298-304.

- (88) Pelletier, R., Krasilnikova, M. M., Samadashwily, G. M., Lahue, R., and Mirkin, S. M. (2003) Replication and expansion of trinucleotide repeats in yeast. *Mol Cell Biol* 23, 1349-1357.
- (89) Krasilnikova, M. M., and Mirkin, S. M. (2004) Replication stalling at Friedreich's ataxia (GAA)(n) repeats in vivo. *Molecular and Cellular Biology* 24, 2286-2295.
- (90) Kang, S. M., Ohshima, K., Shimizu, M., Amirhaeri, S., and Wells, R. D. (1995) Pausing of DNA-Synthesis in-Vitro at Specific Loci in Ctg and Cgg Triplet Repeats from Human Hereditary-Disease Genes. *J Biol Chem* 270, 27014-27021.
- (91) Usdin, K., and Woodford, K. J. (1995) Cgg Repeats Associated with DNA Instability and Chromosome Fragility Form Structures That Block DNA-Synthesis in-Vitro. *Nucleic Acids Res* 23, 4202-4209.
- (92) Kang, S., Jaworski, A., Ohshima, K., and Wells, R. D. (1995) Expansion and deletion of CTG repeats from human disease genes are determined by the direction of replication in *E. coli*. *Nat Genet* 10, 213-218.
- (93) Miret, J. J., Pessoa-Brandao, L., and Lahue, R. S. (1998) Orientation-dependent and sequence-specific expansions of CTG/CAG trinucleotide repeats in *Saccharomyces cerevisiae*. *Proceedings of the National Academy of Sciences* 95, 12438-12443.
- (94) Lahue, R. S., Au, K. G., and Modrich, P. (1989) DNA mismatch correction in a defined system. *Science* 245, 160-164.
- (95) Manley, K., Shirley, T. L., Flaherty, L., and Messer, A. (1999) Msh2 deficiency prevents in vivo somatic instability of the CAG repeat in Huntington disease transgenic mice. *Nat Genet* 23, 471-473.

- (96) Savouret, C., Garcia-Cordier, C., Megret, J., te Riele, H., Junien, C., and Gourdon, G. (2004) MSH2-dependent germinal CTG repeat expansions are produced continuously in spermatogonia from DM1 transgenic mice. *Mol Cell Biol* 24, 629-637.
- (97) Owen, B. A., Yang, Z., Lai, M., Gajec, M., Badger, J. D., 2nd, Hayes, J. J., Edelman, W., Kucherlapati, R., Wilson, T. M., and McMurray, C. T. (2005) (CAG)(n)-hairpin DNA binds to Msh2-Msh3 and changes properties of mismatch recognition. *Nat Struct Mol Biol* 12, 663-670.
- (98) Kovtun, I. V., Liu, Y., Bjoras, M., Klungland, A., Wilson, S. H., and McMurray, C. T. (2007) OGG1 initiates age-dependent CAG trinucleotide expansion in somatic cells. *Nature* 447, 447-U442.
- (99) Jarem, D. A., Wilson, N. R., Schermerhorn, K. M., and Delaney, S. (2011) Incidence and persistence of 8-oxo-7,8-dihydroguanine within a hairpin intermediate exacerbates a toxic oxidation cycle associated with trinucleotide repeat expansion. *DNA Repair* 10, 887-896.
- (100) Wells, R. D., Dere, R., Hebert, M. L., Napierala, M., and Son, L. S. (2005) Advances in mechanisms of genetic instability related to hereditary neurological diseases. *Nucleic Acids Res* 33, 3785-3798.
- (101) McGhee, J. D., and Felsenfeld, G. (1980) Nucleosome structure. *Annu Rev Biochem* 49, 1115-1156.
- (102) Luger, K., Mader, A. W., Richmond, R. K., Sargent, D. F., and Richmond, T. J. (1997) Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature* 389, 251-260.
- (103) Bowman, G. D., and Poirier, M. G. (2015) Post-translational modifications of histones that influence nucleosome dynamics. *Chem Rev* 115, 2274-2295.

- (104) Strahl, B. D., and Allis, C. D. (2000) The language of covalent histone modifications. *Nature* 403, 41-45.
- (105) Bannister, A. J., and Kouzarides, T. (2011) Regulation of chromatin by histone modifications. *Cell Res* 21, 381-395.
- (106) Johnson, C. M. (2013) Differential scanning calorimetry as a tool for protein folding and stability. *Arch Biochem Biophys* 531, 100-109.
- (107) Clas, S.-D., Dalton, C. R., and Hancock, B. C. (1999) Differential scanning calorimetry: applications in drug development. *Pharmaceutical Science & Technology Today* 2, 311-320.
- (108) Chowdhry, B. (1989) Differential scanning calorimetry: applications in biotechnology. *Trends in Biotechnology* 7, 11-18.
- (109) Anson, M. L. (1945) Protein Denaturation and the Properties of Protein Groups. 2, 361-386.
- (110) Jackson, S. E. (1998) How do small single-domain proteins fold? *Folding and Design* 3, R81-R91.
- (111) Mayne, L., and Englander, S. W. (2000) Two-state vs. multistate protein unfolding studied by optical melting and hydrogen exchange. *Protein Sci* 9, 1873-1877.

**Chapter 2: Klenow Fragment Discriminates Against the
Incorporation of the Hyperoxidized dGTP Lesion
Spiroiminodihydantoin into DNA[†]**

[†]Adapted and Modified From:

Huang, J., Yennie, C. J., and Delaney, S. (2015) Klenow Fragment
Discriminates Against the Incorporation of the Hyperoxidized dGTP Lesion
Spiroiminodihydantoin into DNA. *Chem. Res. Toxicol.* 28, 2325-2333

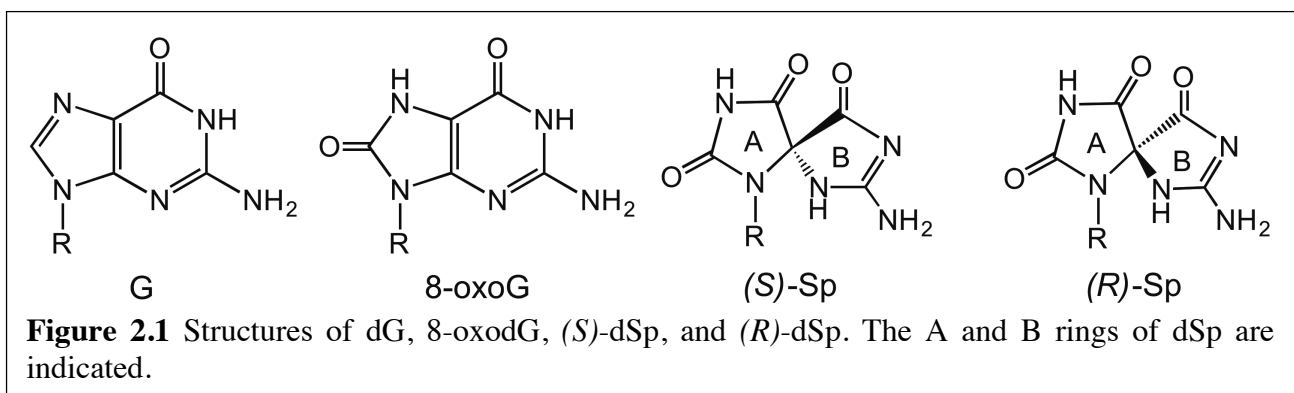
2.1 Abstract

Defining the biological consequences of oxidative DNA damage remains an important and ongoing area of investigation. At the foundation of understanding the repercussions of such damage is a molecular-level description of the action of DNA-processing enzymes, such as polymerases. In this work we focus on a secondary, or hyperoxidized, oxidative lesion of dG which is formed by oxidation of the primary oxidative lesion 2'-deoxy-8-oxo-7,8-dihydroguanosine (8-oxodG). In particular, we examine incorporation into DNA of the diastereomers of the hyperoxidized guanosine triphosphate lesion spiroiminodihydantoin-2'-deoxynucleoside-5'-triphosphate (dSpTP). Using kinetic parameters we describe the ability of the Klenow fragment of *E. coli* DNA polymerase I lacking 3'→5' exonuclease activity (KF^-) to utilize (*S*)-dSpTP and (*R*)-dSpTP as building blocks during replication. We find that both diastereomers act as covert lesions, similar to a Trojan horse: KF^- incorporates the lesion dNTP opposite dC, which is a non-mutagenic event, however, during the subsequent replication it is known that dSp is nearly 100% mutagenic. Nevertheless, using k_{pol}/K_d to define the nucleotide incorporation specificity we find that the extent of oxidation of the dGTP-derived lesion correlates with its ability to be incorporated into DNA. KF^- has the highest specificity for incorporation of dGTP opposite dC. The selection factors for incorporating 8-oxodGTP, (*S*)-dSpTP, and (*R*)-dSpTP are 1,700-, 64,000- and 850,000-fold lower respectively. Thus, KF^- is rigorous in its discrimination against incorporation of the hyperoxidized lesion and these results suggest that the specificity of cellular polymerases provides an effective mechanism to avoid incorporating dSpTP lesions into DNA from the nucleotide pool.

2.2 Introduction

Genomic DNA is exposed to a variety of endogenous and exogenous reactive oxygen species (ROS).^{1,2} Due to its low redox potential,³ dG is the major target for ROS and a prototypic and well-studied oxidation product of dG is 2'-deoxy-8-oxo-7,8-dihydroguanosine (8-oxodG) (Figure 2.1).⁴⁻⁶ Indeed, *in vivo* 8-oxodG is found at steady-state levels of 0.3-4.2 per 10⁶ dG.⁷ Studies performed *in vitro* have shown that 8-oxodG has a high miscoding frequency because it can base pair with both dC and dA, with the latter leading to G→T transversion mutations.^{8,9} Despite its high miscoding frequency *in vitro*, 8-oxodG has a low mutation frequency *in vivo* yielding less than 10% G→T transversions.⁹⁻¹¹ This low mutagenicity *in vivo* derives from the presence of an extensive repair system that counters the genetic effects of 8-oxodG.¹²

In addition to the mutagenic potential of 8-oxodG, the lesion itself is chemically labile towards further oxidation.¹³ Several hyperoxidized lesions have been identified, including 2'-deoxy-spiroiminodihydantoin (dSp), which exists as a pair of diastereomers (Figure 2.1).¹⁴⁻²⁰



Furthermore, the dSp lesion has been detected in genomic DNA from bacteria and mammalian cells. Hailer and co-workers detected the dSp lesion in *E. coli* treated with chromate,²¹ and more recently, Mangerich and co-workers identified dSp in *Helicobacter hepaticus*-infected mice, which develop inflammation-mediated carcinogenesis at levels of 1-5 per 10⁸ nucleotides in genomic DNA.²² Several DNA glycosylases have been found to excise the dSp lesion from DNA

in vitro and initiate the base excision repair (BER) process. The bacterial glycosylases FPG, Nei, and Nth remove dSp with differing preferences for the opposing base pair partner.²³⁻²⁵ The dSp lesion is not a substrate for the human enzyme OGG1 which is responsible for removing 8-oxodG, but the lesion is a substrate for yeast OGG1 and OGG2.²⁶ Finally, eukaryotic²⁷⁻³⁰ and viral³¹ NEIL1 glycosylases can excise dSp from double-stranded DNA with some preference for the opposing base pair partner, and can also remove dSp from single-stranded DNA, bubble, and bulge structures. More recent results suggest that dSp can also be removed from DNA by nucleotide excision repair.³² If dSp is not removed prior to replication, in contrast to the mildly mutagenic 8-oxodG, the hyperoxidized dSp lesion is nearly 100% mutagenic *in vitro*^{33, 34} and *in vivo*,³⁵⁻³⁷ causing both G→T and G→C transversions.

As the primary and secondary oxidation products of dG, the mutagenicity of both 8-oxodG and dSp have been well-studied when the lesion is present in DNA. However, in order to understand the full mutagenic potential of a lesion we must also understand the extent to which polymerases utilize the oxidized lesions when they are part of the nucleotide pool. Notably, it has been shown that a nucleobase is more susceptible to oxidation when it is part of a free nucleotide than when incorporated in duplex DNA. For instance, dATP in the nucleotide pool is 67-fold more easily oxidized than dA in duplex DNA.³⁸ Although the preference is less pronounced for dGTP, it is still 9-fold easier to oxidize in the nucleotide pool relative to dG in duplex DNA.³⁸ Therefore, it is important to consider the nucleotide pool as a source of oxidative damage. Indeed, it is known that 8-oxodGTP can be incorporated into DNA by several bacterial and mammalian polymerases.³⁹⁻⁴³ However, it remains unknown if polymerases can utilize dSpTP as a building block during DNA replication.

Here we determine the extent to which each diastereomer of dSpTP is a substrate for the

Klenow Fragment of *E. coli* DNA polymerase I lacking 3'→5' exonuclease activity (KF⁻). We establish that KF⁻ can incorporate each lesion into DNA, with a preference for (*S*)-dSpTP over (*R*)-dSpTP. We also define the kinetic parameters for incorporation of each lesion and discuss the incorporation efficiency relative to canonical dGTP and the primary oxidation product 8-oxodGTP.

2.3 Experimental Procedure

2.3.1 Materials

Klenow Fragment 3'→5' exo⁻ (KF⁻), calf intestinal alkaline phosphatase, and T4 polynucleotide kinase were purchased from New England Biolabs (Ipswich, MA). The canonical dNTPs were from Thermo Scientific (Waltham, MA) and 8-oxodGTP was from TriLink BioTechnologies (San Diego, CA).

2.3.2 Oligonucleotide Synthesis and Purification

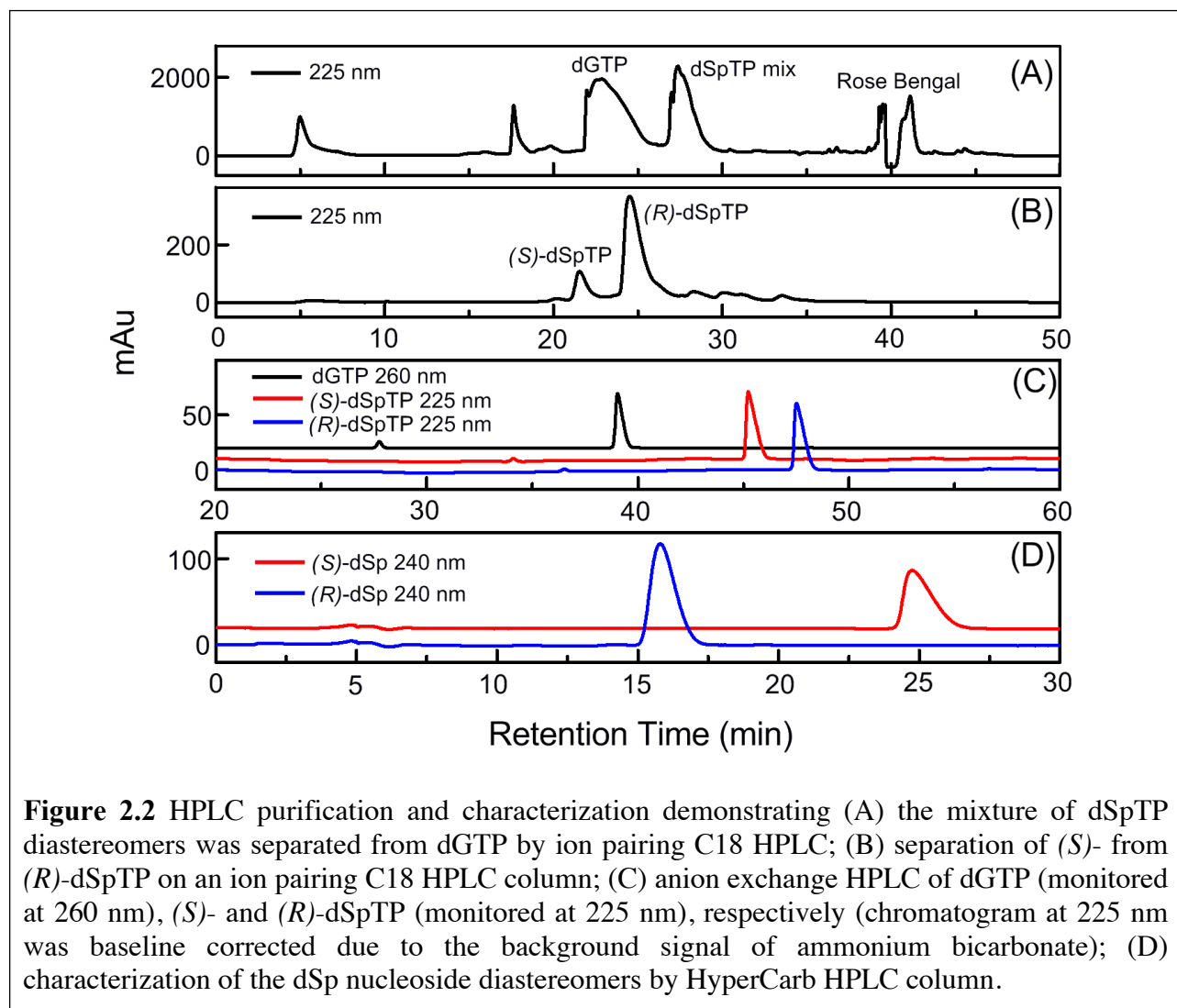
Oligonucleotides were synthesized and purified according to methods published previously by our laboratory.⁴⁴ Oligonucleotide concentrations were determined at 90 °C using molar extinction coefficients as determined by nearest-neighbor theory for single-stranded DNA^{45, 46} on a Beckman Coulter DU800 UV-visible spectrophotometer equipped with a peltier thermoelectric device. The identity of the oligonucleotides was confirmed by ESI-MS.

2.3.3 Synthesis, Purification, and Absolute Configuration of Spiroiminodihydroantoin-2'-deoxynucleoside-5'-triphosphate

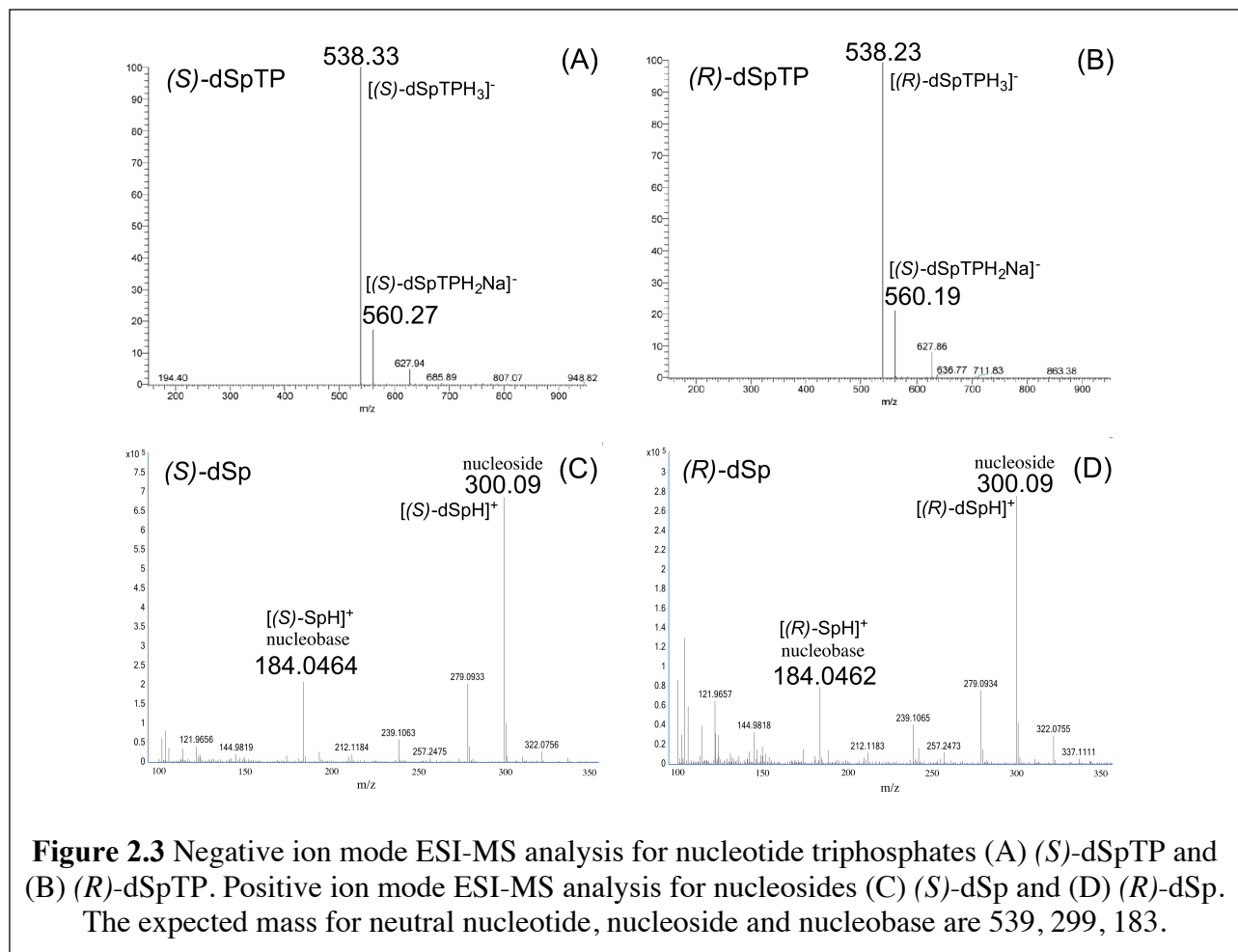
The hyperoxidized nucleotide triphosphate dSpTP was synthesized according to literature

procedures from dGTP.⁴⁷ The resulting solution was dried *in vacuo*, resuspended in a small amount of deionized water, and purified by ion-pairing HPLC on a Varian Microsorb C18 column (250 · 4.6 mm) (Agilent Technologies, Inc., Santa Clara, CA) using acetonitrile (mobile phase A) and 10 mM tetrabutylammonium hydroxide, 10 mM monopotassium phosphate, 0.25% methanol, pH = 7.0 (mobile phase B) as the mobile phases. As the gradient mobile phase A increased from 5 to 65% over 40 min at a flow rate of 1 mL/min. The dSpTP diastereomers were collected as a mixture and dried *in vacuo*; Figure 2.2A demonstrates that the mixture of dSpTP diastereomers was successfully separated from the dGTP starting material. The sample was resuspended in deionized water and the diastereomers were separated on the same HPLC column as above using mobile phase B (same as above) and mobile phase C (2.8 mM tetrabutylammonium hydroxide, 100 mM monopotassium phosphate, 30% methanol, pH = 5.5); as the gradient mobile phase C was increased from 50 to 70% over 40 min at flow rate 0.9 mL/min (Figure 2.2B). Removal of the ion-pairing reagent from the separated diastereomers was achieved by multiple rounds of purification on a Dionex DNAPac PA-100 anion exchange HPLC column (250 · 4.6 mm) (Thermo Scientific, Waltham, MA) using mobile phase D (25 mM ammonium bicarbonate, pH=8.5) and mobile phase E (800 mM ammonium bicarbonate, pH=8.5); as the gradient mobile phase E was increased from 0 to 50% over 15 min at a flow rate of 1 mL/min. Multiple rounds of anion exchange HPLC were performed to the point that the experimentally determined rate of nucleotide incorporation did not change. Notably, the diastereomers of dSpTP elute in the same order from the C18 and anion exchange HPLC columns. The volatile ammonium bicarbonate was removed by repetitive resuspension in deionized water and drying *in vacuo*. For all the HPLC analyses described above, absorbance was monitored at both 225 nm and 260 nm. The purified dSpTP diastereomers were characterized along with dGTP by anion exchange HPLC using mobile phase D (25 mM

ammonium bicarbonate, pH=8.5) and mobile phase E (800 mM ammonium bicarbonate, pH=8.5); as the gradient mobile phase E was increased from 0 to 50% over 60 min at a flow rate of 1 mL/min (Figure 2.2C). The minor peak observed in panel C for dGTP, (*S*)-dSpTP, and (*R*)-dSpTP corresponds to the dNDP hydrolysis product which occurs during freeze/thaw cycles of dNTPs. Importantly, the dNDP version is not a substrate for polymerase and would not influence our kinetic results. The concentration of each diastereomer of dSpTP was determined by UV absorbance based on the extinction coefficient $\epsilon_{230\text{nm}} = 4,900 \text{ M}^{-1} \text{ cm}^{-1}$.¹⁴ The molecular weights were verified by ESI-MS (Figure 2.3A, B).



The absolute configuration of the diastereomers of dSpTP was then assigned based on literature reports.^{48, 49} Each diastereomer of dSpTP (0.2 μmol) was resuspended in 1X Cutsmart buffer (20 mM Tris acetate, 50 mM potassium acetate, 10 mM magnesium acetate, 0.1 mg/mL BSA, pH 7.9) (New England Biolabs, Ipswich, MA) and treated with 20 U calf intestinal alkaline phosphatase at 37 °C for 40 h. The phosphatase was removed by passing the solution through a 3,000 MWCO centrifugal filter (Sartorius Corp., Bohemia, NY). The resulting dSp nucleosides were then analyzed on a Thermo Scientific HyperCarb column (150 \times 4.6 mm) (Thermo Scientific, Waltham, MA) with an isocratic flow of deionized water containing 0.1 % acetic acid at 1 mL/min; the absorbance was monitored at 240 nm (Figure 2.2D). The dSp nucleosides were then analyzed by ESI-MS (Figure 2.3C, D).



Consistent with literature reports,⁴⁹ removal of the triphosphate group switches the order of elution on the HyperCarb column relative to the C18 and anion exchange columns. Based on literature precedent,⁴⁹ we assigned the absolute configurations of (*S*)-dSpTP (triphosphate elutes first from C18 and anion exchange columns, nucleoside elutes second from HyperCarb column) and (*R*)-dSpTP (triphosphate elutes second from C18 and anion exchange columns, nucleoside elutes first from HyperCarb column).

2.3.4 Preparation of DNA Primer/Template Assemblies

The primer strand was radiolabeled at the 5'-end with ³²P using T4 polynucleotide kinase and annealed to the template in a 1:1.5 ratio of primer:template in a buffer containing 10 mM Tris-HCl, 50 mM NaCl, 10 mM MgCl₂, and 1 mM DTT, pH 7.9 by heating the solution at 90 °C for 5 min and allowing it to slowly return to room temperature over approximately 2.5 h.

2.3.5 Qualitative Single Nucleotide Triphosphate Incorporation Reactions

Primer/template DNA was incubated with KF⁻ for 5 min on ice where the final sample contained 5 nM primer/template DNA, 25 nM KF⁻, 10 mM Tris-HCl, 50 mM NaCl, 10 mM MgCl₂, 1 mM DTT and 0.04 mg/mL BSA, pH 7.9. The incorporation reaction was initiated by addition of nucleotide triphosphate to yield a final concentration of 20 nM dGTP, 5 μM 8-oxodGTP, 100 μM (*S*)-dSpTP, or 100 μM (*R*)-dSpTP and the samples were incubated for 15 min at room temperature. The reactions were quenched by addition of gel loading buffer (80% formamide, 1 mg/mL xylene cyanol, 1 mg/ml bromophenol blue, and 10 mM EDTA), electrophoresed through a 16% denaturing PAGE gel, and imaged by phosphorimager.

2.3.6 Primer Extension after the Incorporation of Sp Nucleotide Triphosphate

To examine whether primer extension will continue after the incorporation of the dSpTP nucleotide triphosphates, DNA with template X = C was used. In addition to 100 μ M of dSpTP, 20 nM of both dATP and dCTP were added to the reaction while the rest of conditions are the same as described above for the qualitative single nucleotide incorporation experiment.

2.3.7 Single-Turnover Kinetics Reactions

Reactions were performed at 25 °C using a Rapid Quench Flow apparatus (RQF-3) from KinTek Corporation (Austin, TX). A solution consisting of 10 nM primer/template DNA, 50 nM KF^- , 10 mM Tris-HCl, 50 mM NaCl, 10 mM MgCl_2 , 1 mM DTT, 0.08 mg/mL BSA, pH 7.9 was incubated for 5 min on ice. The incorporation reaction was initiated by adding various concentrations of each nucleotide triphosphate (listed in each figure legend) in buffer containing 10 mM Tris-HCl, 50 mM NaCl, 10 mM MgCl_2 , and 1 mM DTT, pH 7.9. For each reaction in the RQF, 15 μ L DNA· KF^- and 15 μ L dNTP were loaded into the reaction loops, rapidly mixed and incubated for the designated time before being quenched by 85 μ L of 100 mM EDTA. The final reaction concentrations were 5 nM primer/template DNA, 25 nM KF^- , dNTP (various concentrations), 10 mM Tris-HCl, 50 mM NaCl, 10 mM MgCl_2 , 1 mM DTT, 0.04 mg/mL BSA, pH 7.9. Gel loading buffer was added and the samples were analyzed by 16% denaturing PAGE gel and imaged by phosphorimager.

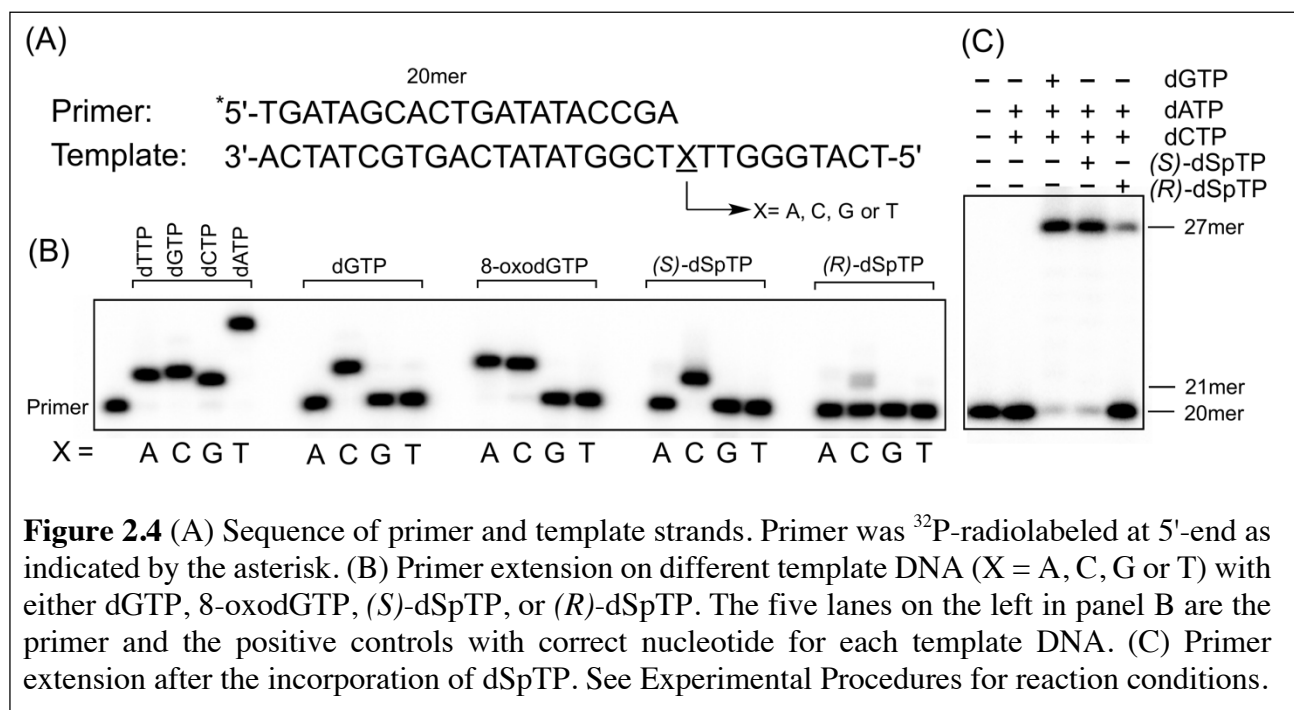
2.3.8 Data Analysis

KinTek Explorer software (KinTek Corporation, Austin, TX) was used to globally fit the kinetic data.^{50, 51} The data were iteratively fit until a best fit was obtained according to a specific

enzyme model. Because binding of a canonical nucleotide triphosphate to a polymerase is diffusion limited, k_1 was fixed at $150 \mu\text{M}^{-1} \text{s}^{-1}$ for dGTP.⁵² However, for the oxidized and hyperoxidized nucleotide triphosphates 8-oxodGTP and dSpTP, k_1 was fixed at $10 \mu\text{M}^{-1} \text{s}^{-1}$ due to their reduced binding.⁵¹ The lower and upper limits associated with each kinetic parameter were defined using the FitSpace function of KinTek Explorer. The best fit curves and experimental data were exported from KinTek Explorer and plotted in KaleidaGraph (Synergy software, Reading, PA).

2.4 Results and Discussion

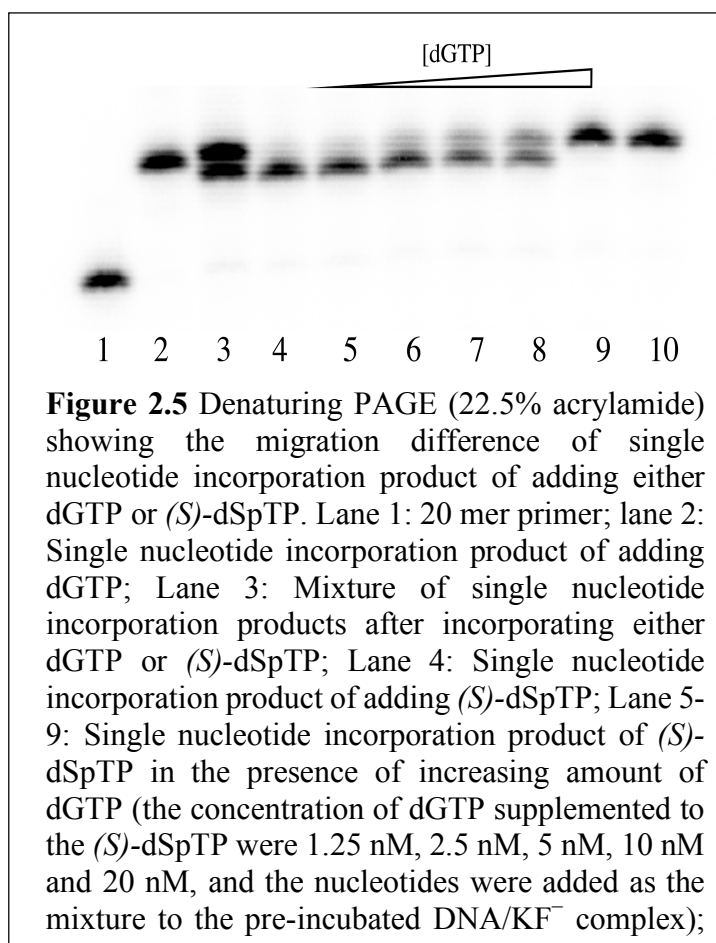
2.4.1 Qualitative Analysis of Incorporation of Spiroiminodihydantoin Nucleotide Triphosphate by Klenow Fragment



In order to study incorporation of the hyperoxidized nucleotide triphosphate dSpTP into DNA, templates were designed with either X = dA, dC, dG, or dT at the site of incorporation (Figure 2.4A). The diastereomers of dSpTP were resolved and examined separately. Primer

extension reactions were catalyzed by KF^- and were initiated by the addition of either dGTP, 8-oxodGTP, (*S*)-dSpTP, or (*R*)-dSpTP. The results are shown in Figure 2.4B.

As expected, dGTP is incorporated exclusively opposite dC, and 8-oxodGTP is incorporated opposite both dC and dA.⁸ Both diastereomers of dSpTP are incorporated opposite dC, with more incorporation observed for (*S*)-dSpTP relative to (*R*)-dSpTP. Notably, in order to observe sufficient amounts of the extended primer product, the concentration of 8-oxodGTP was 250-fold greater than dGTP, and the concentration of the dSpTP diastereomers was 5,000-fold greater than dGTP. The experiments were also performed with dSpTP that was prepared using a different synthetic strategy (using 8-oxodGTP as the starting material⁴⁷ rather than dGTP), and the same results were obtained (data not shown). Obtaining the same results with these different preparations of dSpTP indicates that these polymerase incorporation results are not due to trace contaminants of the dGTP starting material that may remain after synthesis and purification, but rather to dSpTP itself. That the product observed in Figure 2.4B corresponds to incorporation of (*S*)- and (*R*)-dSpTP rather than potential contaminating dGTP starting material is further supported by the difference in the



migration between the single nucleotide incorporation products when dGTP or our preparation of dSpTP are added to the primer (Figure 2.5).

Having observed that both (*S*)-dSpTP and (*R*)-dSpTP can be incorporated into DNA from the nucleotide pool, we next examined whether the polymerase can extend beyond dSpTP or if incorporation of the hyperoxidized nucleotide would terminate replication. To avoid competition between dSpTP and dGTP, the dGTP was excluded from the reactions. Both dATP and dCTP were included and under these experimental conditions the full extension product would be a 27-mer. As shown in Figure 2.4C, KF⁻ can extend past both diastereomers of dSpTP to yield the 27-mer product, suggesting that if incorporated from the nucleotide pool the lesions are not absolute blocks to replication. It is noteworthy that these qualitative experiments were performed using an excess of polymerase, and that under conditions limiting in polymerase stalling at the lesion site could occur.

In previous work, it was shown that when dSp is located in the template strand, KF⁻ incorporates dATP and dGTP opposite the lesion both *in vitro* and *in vivo*.³³⁻³⁷ In fact, the correct base dCTP is almost never incorporated opposite dSp, meaning that the lesion is nearly 100% mutagenic if formed in genomic DNA. Based on these prior observations, we expected the dSpTP diastereomers to be incorporated into templates that contain X = dA and dG. Consistent with this logic, 8-oxodG is paired with dA and dC regardless of whether the lesion is in the template or is incorporated from the nucleotide pool.^{8, 33, 40, 42} Our observation that KF⁻ only incorporates the dSpTP diastereomers opposite of dC in the template provides an intriguing example where a polymerase exhibits a different preference based on whether the lesion is in the template or in the nucleotide pool.

It is of note that because the hyperoxidized lesion is a dG-derived lesion, the incorporation of dSpTP opposite dC does not immediately represent a mutagenic event. However, in a subsequent round of replication the hyperoxidized lesion will serve as the template and in this context is known to be nearly 100% mutagenic. Thus, one could liken the dNTP form of the hyperoxidized lesion to a Trojan horse; the dNTP is incorporated into DNA by polymerase but will not exhibit mutagenicity until a G → C or G → T mutation occurs in the next round of replication.

It is possible to rationalize how a polymerase could use different base pairing for a lesion, depending on whether the lesion serves as the template or is incorporated from the nucleotide pool. Prior to making a covalent linkage between the primer and incoming nucleotide, a polymerase probes the template nucleobase for favorable electrostatic interactions. It has been proposed that for KF^- the electrostatic profile is a significant determinant in nucleotide selectivity.⁵³ Interestingly, the B-ring of the lowest energy tautomer of the diastereomers of dSpTP⁴⁹ can form two of the same hydrogen bonds as the Watson-Crick face of dGTP (using a lone pair on the carbonyl and a hydrogen of the exocyclic amine) suggesting a means by which the hyperoxidized lesion pairs with dC. Other tautomers of dSpTP (i.e., the imino tautomer or the amino unconjugated tautomer⁴⁹), although energetically less favorable, can form three hydrogen bonds with dC. Furthermore, a nucleotide triphosphate in solution has more conformational freedom than a nucleobase in a template strand because it is not confined by base stacking, base pairing, and covalent attachment to the sugar-phosphate backbone. This conformational freedom could allow incoming dSpTP to adopt the *anti* conformation and present the B-ring to base pair with dC in the template. However, characterizations of dSp have shown that the A- and B-rings are perpendicular to one another.⁵⁴ Due to this structural constraint, when the lesion is located in the template, the

A-ring of dSp could be presented for base pairing with an incoming nucleotide, which would explain the lack of pairing with dCTP when the lesion is in the template. Indeed, a recent crystal structure of (*S*)-dSp as the templating nucleobase in the active site of DNA polymerase β (pol β) reveals that the lesion is in the *syn* conformation and uses the A-ring as the base pairing face.⁵⁴

While the information is not yet available for Sp, crystal structures of primer/template-polymerase-dNTP ternary complexes where dSp is either the templating base or dSpTP is the incoming nucleotide would greatly inform our results. Previous ternary structures of pol β bound to DNA containing 8-oxodG as the templating base or as the incoming 8-oxodGTP revealed the molecular basis for different base pairing. When 8-oxodG served as the templating base, the lesion was observed while base pairing with an incoming dCTP. Both 8-oxodG and dCTP were in the *anti* conformation and Watson-Crick base pairing was observed.⁵⁵ In the ternary structure where 8-oxodGTP was the incoming nucleotide the templating base was dA.⁵⁶ While dA remained *anti* the 8-oxodGTP was in the *syn* conformation. The *syn* conformation was stabilized by Hoogsteen hydrogen-bonding with the templating dA. In other work with human DNA polymerase η , which replicates a templating 8-oxodG in a nearly error-free manner, ternary complexes revealed that an Arg residue from a finger domain prevents formation of the 8-oxodG:A mispair.⁵⁷ Similar structural studies would reveal the molecular basis for why a templating dSp base pairs with dATP and dGTP, whereas the nucleotide version dSpTP base pairs with a templating dC.

2.4.2 Kinetic Analyses of Nucleotide Triphosphate Incorporation

To evaluate the biological relevance of incorporation of dSpTP into DNA from the nucleotide pool, we performed experiments to compare the kinetic parameters k_{pol} and K_d of the hyperoxidized nucleotide to dGTP and 8-oxodGTP. To avoid the complications that arise from

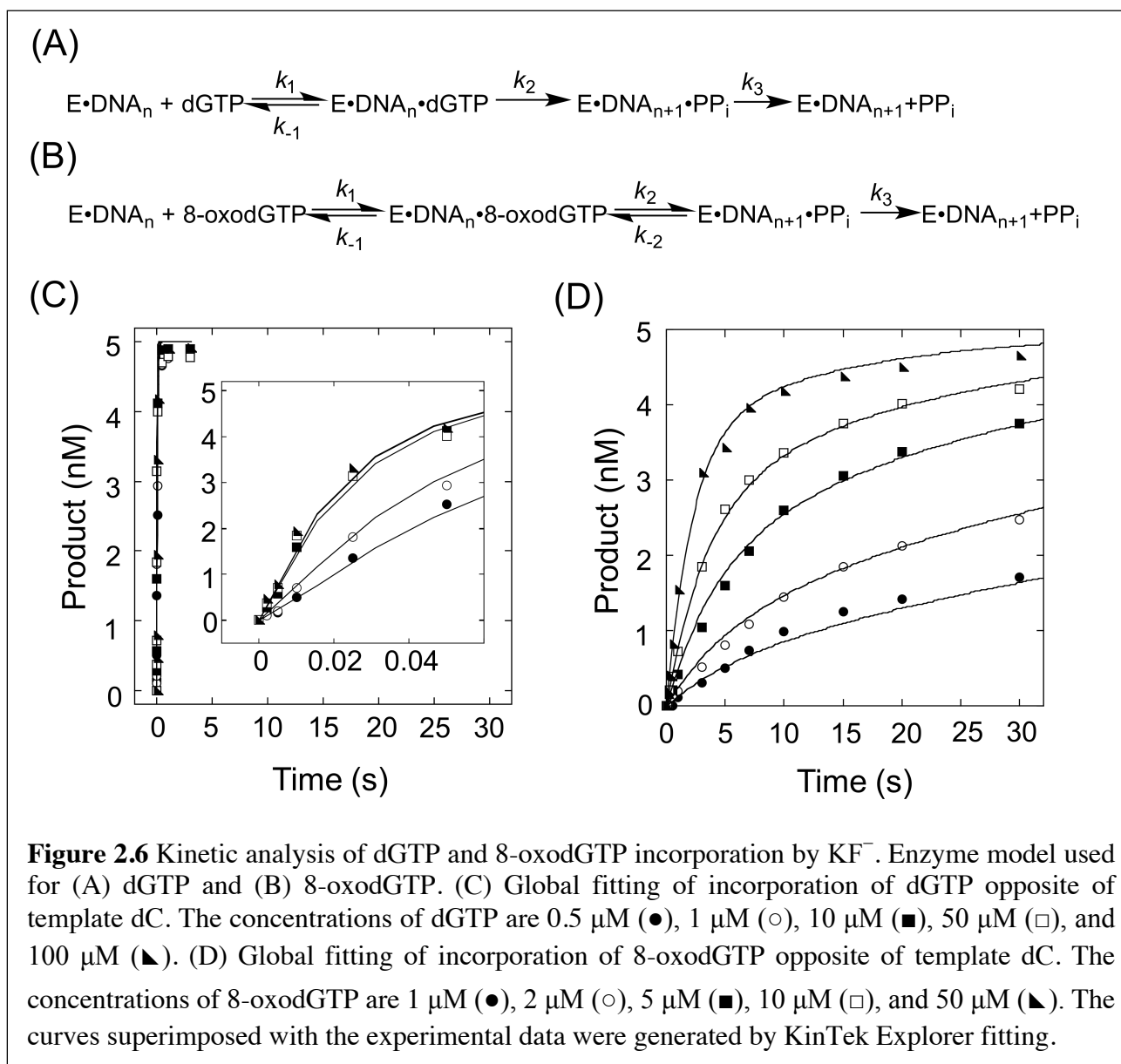
multiple turnovers of a processive polymerase, i.e., the rate being dominated by dissociation of the DNA·KF⁻ complex, the incorporation assays were performed under single-turnover conditions in which [KF⁻] \gg [DNA]. Because single-turnover experiments using DNA polymerases occur rapidly and are too fast to be mixed and quenched manually, we used a rapid quench flow (RQF) apparatus for these experiments. Furthermore, because dGTP, 8-oxodGTP, (*S*)-dSpTP, and (*R*)-dSpTP share the ability to be incorporated opposite C, the kinetic analyses were performed using the template where X = dC.

Single nucleotide incorporation assays were performed using varying concentrations of dNTP in order to examine the concentration dependence of the rate of incorporation and amplitude of product formation. Under these conditions, the concentration dependence of the rate of incorporation reports k_{pol} and K_d . k_{pol} is the rate of the slowest step up to and including the phosphoryl transfer step in which the nucleotide is covalently attached to the primer, and K_d is the ground state nucleotide dissociation constant K_d . The DNA and KF⁻ were pre-incubated prior to initiating the reaction by addition of dNTP. In this manner, the measured K_d is for nucleotide binding to the KF⁻·DNA complex rather than binding of KF⁻ to the DNA. In this report, we use k_{pol}/K_d to define the nucleotide incorporation specificity.⁵⁸

2.4.3 Single-Turnover Kinetic Analysis of dGTP and 8-oxodGTP Incorporation

As the canonical nucleotide to incorporate opposite dC, dGTP was employed as a control for the incorporation experiments. A time course was performed at several dGTP concentrations and the data were fit globally to the model shown in Figure 2.6A. In this model K_d is defined by k_{-1}/k_1 , k_2 is k_{pol} , and k_3 is release of pyrophosphate. For KF⁻ incorporating a canonical dNTP, it is generally accepted that k_{pol} includes a rate limiting conformational change⁵⁹ that controls

nucleotide specificity,⁶⁰ followed by phosphoryl transfer; however, other results suggest that the phosphoryl transfer step might limit the rate.⁶¹ It is also known that for incorporation of canonical nucleotides k_3 is faster than k_2 .⁶² In our analysis, a K_d of 0.8 μM and k_{pol} of 41 s^{-1} define a specificity constant k_{pol}/K_d (or k_2/K_d) of 51 $\mu\text{M}^{-1} \text{s}^{-1}$ for the incorporation of dGTP opposite dC (Table 2.1). It is notable that the final amplitude of 21-mer product is the same at each concentration of dGTP and the slope of the rise converge with increasing concentration (Figure 2.6C).

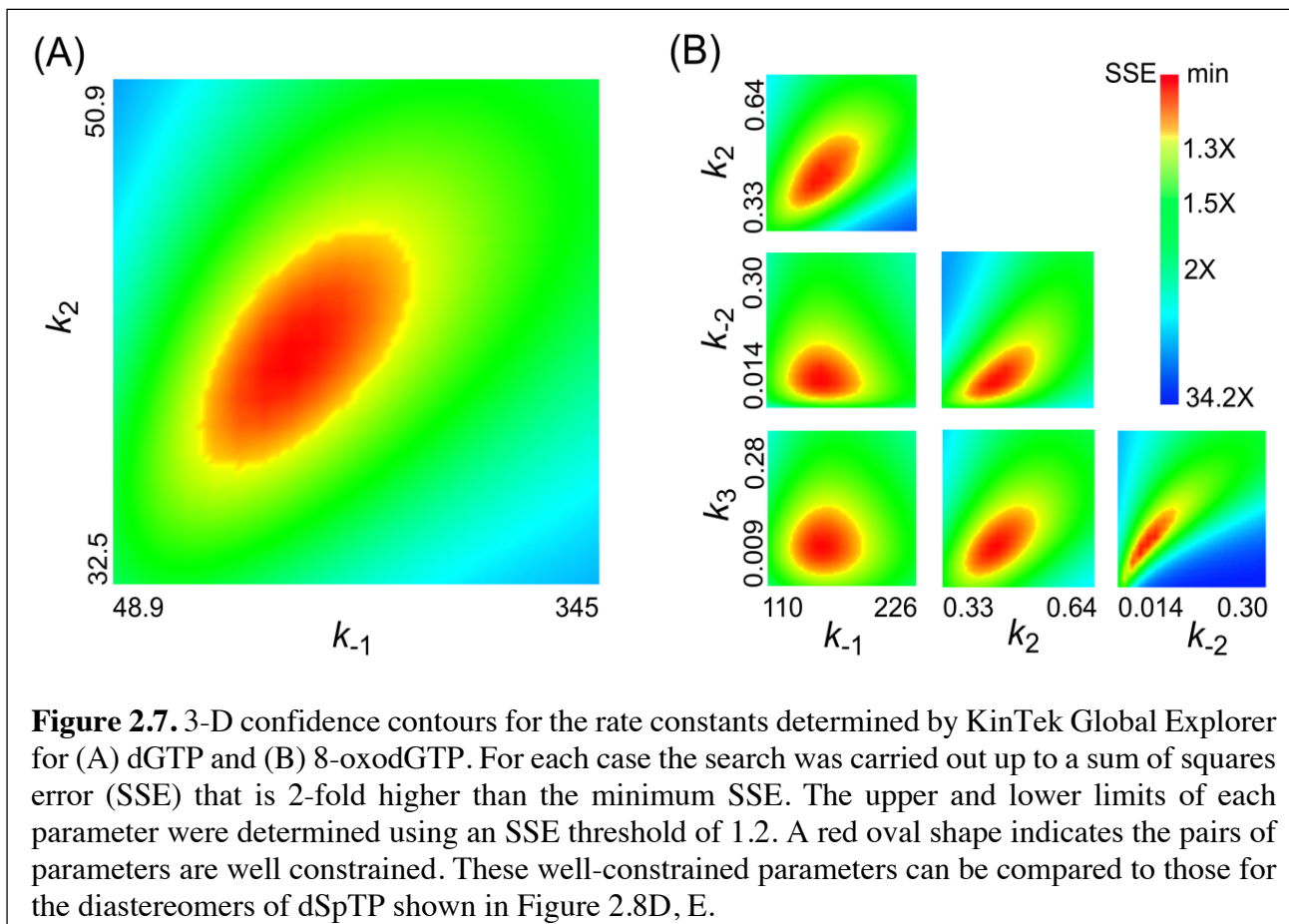


These observations show that k_{pol} is fast and there are no slower steps after incorporation to cause incorporation to be reversible. Both of these conditions are accounted for in the model in Figure 2.6A.

In the case of 8-oxodGTP incorporation opposite of dC, the single-turnover plots at each nucleotide concentration are noticeably biphasic (Figure 2.6D). Additionally, the amplitude is dependent on nucleotide concentration suggesting that after incorporation there is a slow step that allows phosphoryl transfer to be reversible. Indeed, similar results were obtained for the incorporation of 8-oxodGTP opposite dC by the human mitochondrial DNA polymerase.⁴² In accordance with Hanes and co-workers,⁴² we accounted for the differences observed for incorporation of 8-oxodGTP relative to dGTP by fitting the data to the model shown in Figure 2.6B, which allows for reversibility of the phosphoryl transfer. A K_d of 15.2 μM describes nucleotide binding while the incorporation step has a forward rate (k_2) of 0.45 s^{-1} and a reverse rate (k_{-2}) of 0.08 s^{-1} . In contrast to canonical nucleotides, for modified nucleotides it is not known whether k_2 is limited by the conformational change, but it has been shown that when KF incorporates a mismatched nucleotide, the rate of phosphoryl transfer is not limited by the preceding conformational change.⁶³ Pyrophosphate release is indeed slower than incorporation and occurs at a rate of 0.09 s^{-1} , which allows k_2 to be reversible. Therefore, unlike for dGTP, $k_{\text{pol}} \neq k_2$. When taking into account the slow pyrophosphate release, the k_{pol}/K_d value (0.016 $\mu\text{M}^{-1} \text{s}^{-1}$, where $k_{\text{pol}} = k_2 k_3 / (k_{-2} + k_3)$)⁴² is slightly smaller than the k_2/K_d (0.03 $\mu\text{M}^{-1} \text{s}^{-1}$) for 8-oxodGTP. Interestingly, it has been proposed that slow pyrophosphate release is an additional factor that limits incorporation of 8-oxodGTP by mitochondrial polymerase.⁴²

A 3-D confidence contours were provided for both dGTP and 8-oxodGTP kinetic analyses by KinTek software, and the results showed the parameters are well-constrained for both

nucleotides (Figure 2.7). This indicates that the experiments are well set up and the kinetic models accurately represent how polymerase incorporates nucleotides.



2.4.4 Single-Turnover Kinetic Analysis of dSpTP Incorporation

Compared to dGTP and 8-oxodGTP, experiments using the diastereomers of dSpTP required significantly higher nucleotide triphosphate concentrations and extended incubation times to observe a sufficient amount of product for analysis (Figure 2.8B, C). We also observed differences in incorporation between the two dSpTP diastereomers. Specifically, the amount of product accumulated after a particular reaction time for (*S*)-dSpTP is significantly greater than for (*R*)-dSpTP at the same nucleotide concentration, consistent with the results of the

qualitative incorporation presented in Figure 2.4B. Global fitting of the data to the models in Figure 2.6A and 2.6B reveal that the data are described better by a model (Figure 2.8A) that allows reversibility of the incorporation step (data not shown), similar to the results observed for 8-oxodGTP. However, unlike what we observed for incorporation of dGTP or 8-oxodGTP, for the diastereomers of dSpTP the parameters in the kinetic scheme were not well constrained (Figure 2.8D, E and Figure 2.7). As a consequence, we could not obtain individual quantitative values for K_d and k_{pol} . Because k_{pol} is limited by the amount of available DNA·KF⁻·dNTP complex, poor binding of the hyperoxidized nucleotide triphosphate to the DNA·KF⁻ complex likely causes the lack of parameter constraint. Importantly, these data indicate that the K_d for both diastereomers of dSpTP is greater than the nucleotide concentrations used here and these values are at least on the millimolar scale. Even though k_{-1} and k_2 are not constrained and can vary across a wide range for both diastereomers of dSpTP, analysis by KinTek Explorer FitSpace reveals there is a linear correlation between k_2 and k_{-1} (Figure 2.8D, E). Since k_1 was fixed at $10 \mu\text{M}^{-1} \text{s}^{-1}$ during the fitting (see Experimental Procedures), there is also a linear correlation between k_2 and K_d ($K_d = k_{-1}/k_1 = k_{-1}/10$). Therefore we use k_2/K_d to compare the incorporation specificities of the dNTPs used in this study, rather than k_{pol}/K_d . From the slope of FitSpace analysis k_2/k_{-1} (or $k_2/(10 \cdot K_d)$), we defined the incorporation specificity k_2/K_d for (*S*)-dSpTP and (*R*)-dSpTP as $8 \cdot 10^{-4} \mu\text{M}^{-1} \text{s}^{-1}$ and $6 \cdot 10^{-5} \mu\text{M}^{-1} \text{s}^{-1}$, respectively.

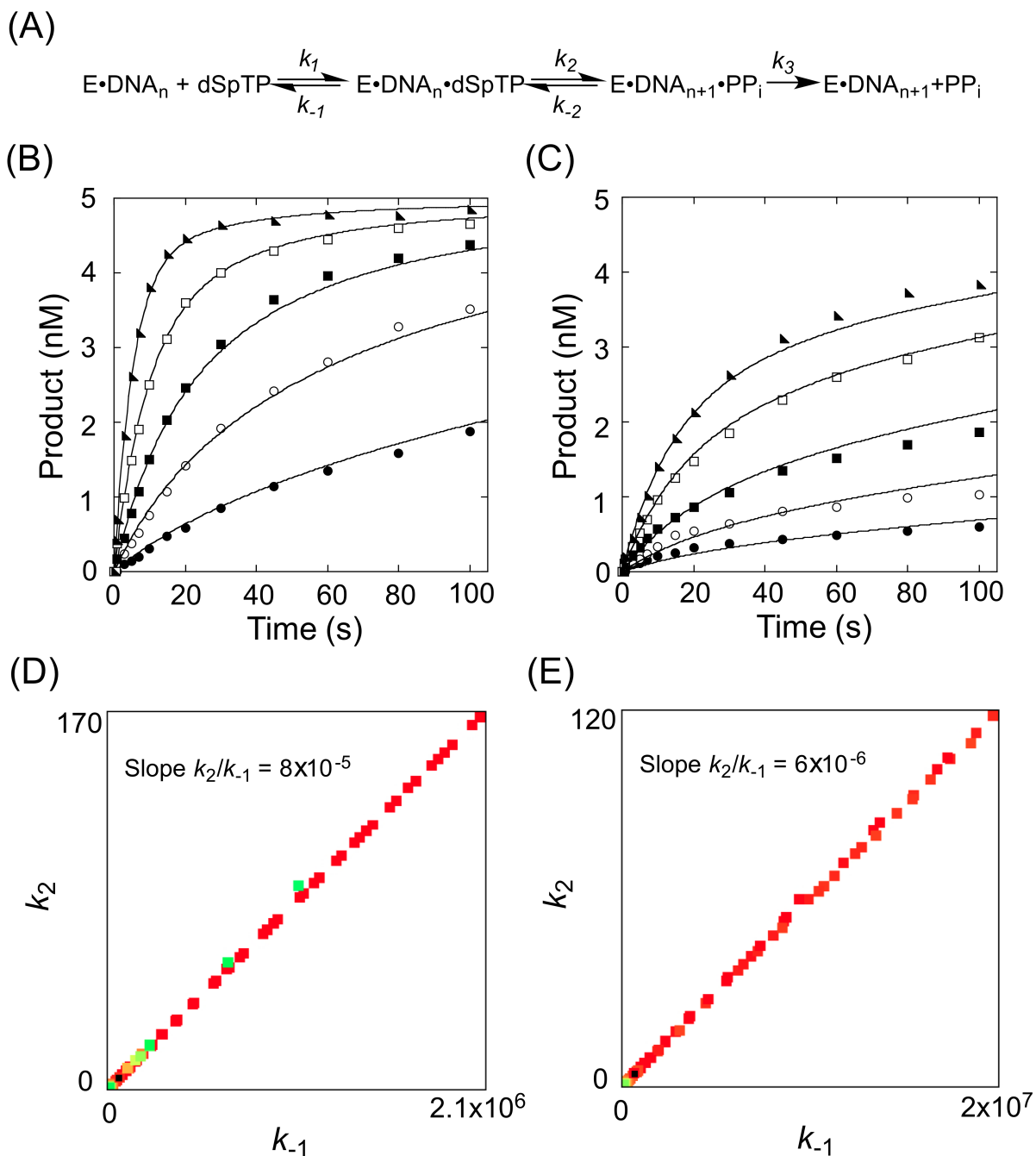


Figure 2.8 Kinetic analysis of (*S*)-dSpTP and (*R*)-dSpTP incorporation by KF^- . (A) Enzyme model used for global fitting. (B) Global fitting of incorporation of (*S*)-dSpTP opposite of template dC. (*S*)-dSpTP concentrations are 10 μM (\bullet), 20 μM (\circ), 50 μM (\blacksquare), 100 μM (\square), and 200 μM (\blacktriangle). (C) Global fitting of incorporation of (*R*)-dSpTP opposite of template dC. (*R*)-dSpTP concentrations are 50 μM (\bullet), 100 μM (\circ), 200 μM (\blacksquare), 400 μM (\square), and 600 μM (\blacktriangle). The curves superimposed with the experimental data were generated by KinTek Explorer fitting. FitSpace analysis by KinTek Explorer showing the results of the initial excursions to map the boundaries of a good fit for (D) (*S*)-dSpTP and (E) (*R*)-dSpTP.

Our kinetic results demonstrate that KF^- exhibits a stereochemical preference and more readily incorporates (*S*)-dSpTP over (*R*)-dSpTP. A general stereochemical preference for the (*S*) diastereomer has been reported previously for several other DNA-processing enzymes. The Burrows laboratory showed that KF^- was more efficient at inserting dATP opposite dSp1 relative to dSp2, where the diastereomers were first described based on their order of elution from an anion exchange HPLC column.³⁴ Subsequent studies revealed that dSp1 and dSp2 are (*S*)-dSp and (*R*)-dSp, respectively.^{48, 49} The Essigmann laboratory similarly showed that pol II and pol IV more efficiently bypass dSp1 *in vivo*⁶⁴ and that the relative amount of $\text{G} \rightarrow \text{C}$ and $\text{G} \rightarrow \text{T}$ transversions can vary for dSp1 and dSp2 depending on the sequence context.³⁵⁻³⁷ Furthermore, nuclease P1⁶⁵ and the BER glycosylase human NEIL1^{28, 30, 32} also have a preference for the dSp1 configuration. Molecular dynamics simulations showed that hNEIL1 makes better contacts with (*S*)-dSp in its binding pocket, justifying the stereochemical preference of the enzyme.⁶⁶

2.4.5 Comparing the Kinetic Parameters for dGTP, 8-oxodGTP, (*S*)-dSpTP, and (*R*)-dSpTP

Our results show that the extent of oxidation of the dGTP-derived lesion is correlated with its ability to be incorporated into DNA. The specificity of incorporation is highest for dGTP opposite dC by KF^- . Compared to dGTP, the selection factors for incorporating 8-oxodGTP, (*S*)-dSpTP, and (*R*)-dSpTP are 1,700-, 64,000- and 850,000-fold lower respectively (Table 2.1). Notably, since pyrophosphate release is slow for 8-oxodGTP, (*S*)-dSpTP, and (*R*)-dSpTP, $k_{\text{pol}}/K_{\text{d}}$ must be smaller than k_2/K_{d} , so these selection factors are lower limits for specificity.

Table 2.1 Kinetic Parameters for the Incorporation of dNTPs Opposite Template dC.

| Nucleotide | K_d (μM) | k_2 (s^{-1}) | k_2/K_d ($\mu\text{M}^{-1}\text{s}^{-1}$) | selection factor ^a |
|--------------------|-------------------------|---------------------------|---|-------------------------------|
| dGTP | 0.8 | 41 | 51 | 1 |
| 8-oxodGTP | 15.2 | 0.45 | 0.030 | 1,700 |
| (<i>S</i>)-dSpTP | 1,700 ^b | 1.4 ^b | $8 \cdot 10^{-4}$ | 64,000 |
| (<i>R</i>)-dSpTP | 7,000 ^b | 0.45 ^b | $6 \cdot 10^{-5}$ | 850,000 |

^aSelection factor is the ratio of the specificity of incorporation (k_2/K_d) of dGTP relative to the dNTP of interest.

^bValues are the lower bound provided by KinTek Explorer global fitting.

In order to provide a graphical representation of the difference in specificity for incorporation of the canonical, oxidized, and hyperoxidized guanine lesions we also analyzed our kinetic data using traditional Michaelis-Menten methods, in which the observed rate (k_{obs}) is plotted against nucleotide triphosphate concentration (Figure 2.9). The slope of the rise can be used to define incorporation specificity k_{pol}/K_d . In fact, the k_{pol}/K_d from this Michaelis-Menten analysis (data not

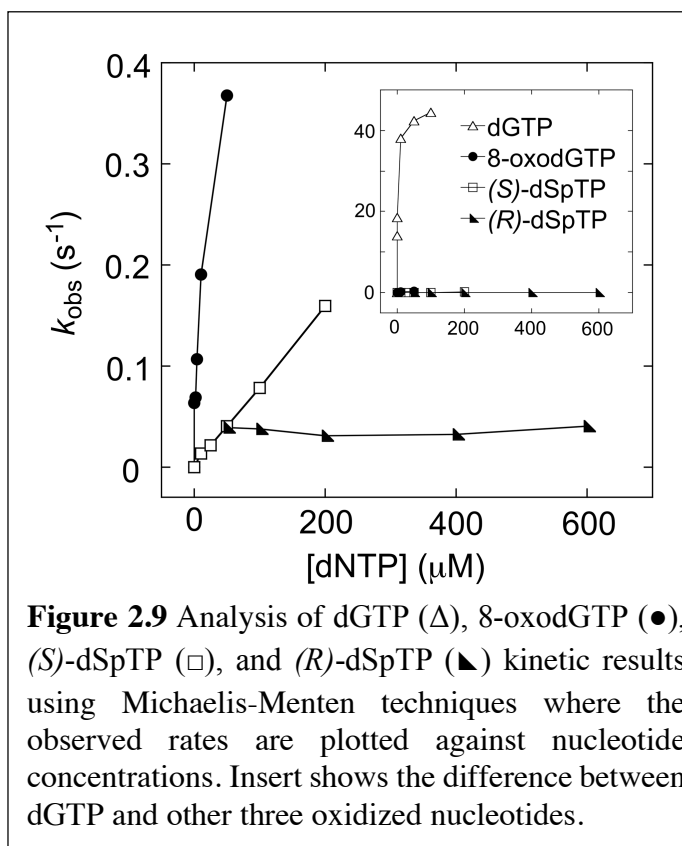


Figure 2.9 Analysis of dGTP (Δ), 8-oxodGTP (\bullet), (*S*)-dSpTP (\square), and (*R*)-dSpTP (\blacktriangle) kinetic results using Michaelis-Menten techniques where the observed rates are plotted against nucleotide concentrations. Insert shows the difference between dGTP and other three oxidized nucleotides.

shown) is comparable to the k_2/K_d obtained by global fitting. Clearly, the incorporation kinetics for dGTP are significantly different from those of oxidized nucleotides (Figure 2.9 inset, open triangles). For the lesion dNTPs, a rise in k_{obs} as a function of dNTP concentrations is observed for 8-oxodGTP and (*S*)-dSpTP, with 8-oxodGTP more steep than (*S*)-dSpTP. However, for (*R*)-dSpTP, k_{obs} was extremely slow across the entire concentration range, and there was no obvious dependence of the rate on nucleotide concentration. These results are entirely consistent with the

quantitative parameters obtained by global fitting.

2.4.6 Biological Considerations

In addition to the kinetic parameters describing incorporation of lesion dNTPs, one must also consider the action of enzymes responsible for cleansing the nucleotide pool. The importance of removing 8-oxodGTP from the nucleotide pool is underscored by the fact that *E. coli* lacking MutT, a phosphatase that converts 8-oxodGTP to 8-oxodGMP,³⁹ have a 100- to 10,000-fold higher mutation rate compared to wild type *E. coli*.⁶⁷ This dramatic increase in mutation rate in the absence of MutT indicates that the nucleotide pool represents a biologically significant source of 8-oxodGTP.^{68, 69} However, it has been demonstrated that the diastereomers of dSpTP are not good substrates for MutT or other known MutT-type nucleotide pool sanitization enzymes in *E. coli*.⁷⁰ Therefore, preventing the incorporation of dSpTP into DNA may depend solely on the ability of a polymerase to discriminate the hyperoxidized nucleotide triphosphate from dGTP. Indeed, consistent with our kinetic results, *in vivo* studies demonstrate that introduction of dSpTP1 or dSpTP2 into *E. coli* cells does not significantly increase the mutation frequency.⁷⁰ This result is in contrast to the high mutagenicity observed in *E. coli* when the Sp lesions are in the DNA template.^{35-37, 64} Our results suggest that the lack of mutagenicity of dSpTP is because the *E. coli* cellular polymerases effectively discriminate the hyperoxidized nucleotides and do not use them as building blocks during replication. It is notable, however, that although it is difficult to force dSpTP into DNA from the nucleotide pool, if incorporation occurs, the lesion is a powerful source of mutations. Indeed, the “Trojan horse” quality of dSpTP makes it an intriguing player in the field of lethal mutagenesis^{71, 72} where a miscoding nucleotide is used as an antiviral agent to accelerate viral mutation rates and drive a viral population to extinction. In order to be effective in this

manner, however, the incorporation efficiency of dSpTP would likely need to be higher than that observed for KF⁻.

2.5 Conclusion

In this work, we used KF⁻ as a model polymerase. Similar incorporation efficiencies have been observed for KF⁻ and some mammalian replicative polymerases incorporating 8-oxodGTP into DNA,⁷³ although mammalian polymerases have a higher preference to incorporate 8-oxodGTP opposite the correct dC instead of the mismatched dA template.⁷⁴ Therefore, we expect the selection factor for mammalian polymerases incorporating the hyperoxidized dNTP to be even more rigorous than KF⁻ and to prevent incorporation of dSpTP from the nucleotide pool.

As a final consideration the *in vivo* concentrations of (*S*)-dSpTP and (*R*)-dSpTP will also influence the extent to which the hyperoxidized lesion can be incorporated into DNA. These values remain to be determined but are most certainly well below the millimolar K_d values suggested by our experiments. Thus, we conclude that while dSpTP can be described as a Trojan horse, the specificity of cellular polymerase provides an effective defense against using the hyperoxidized nucleotide as a building block during DNA replication, and that the dSp lesion is a much more potent mutagen when it is formed in genomic DNA.

2.6 Reference

- (1) Gates, K. S. (2009) An overview of chemical processes that damage cellular DNA: spontaneous hydrolysis, alkylation, and reactions with radicals. *Chem. Res. Toxicol.* 22, 1747-1760.

- (2) Cadet, J., Douki, T., and Ravanat, J. L. (2010) Oxidatively generated base damage to cellular DNA. *Free Radical Biol. Med.* 49, 9-21.
- (3) Steenken, S., and Jovanovic, S. V. (1997) How easily oxidizable is DNA? One-electron reduction potentials of adenosine and guanosine radicals in aqueous solution. *J. Am. Chem. Soc.* 119, 617-618.
- (4) Burrows, C. J., and Muller, J. G. (1998) Oxidative nucleobase modifications leading to strand scission. *Chem. Rev.* 98, 1109-1152.
- (5) Cadet, J. (2003) Oxidative damage to DNA: formation, measurement and biochemical features. *Mutat. Res., Fundam. Mol. Mech. Mutagen.* 531, 5-23.
- (6) Delaney, S., Jarem, D. A., Volle, C. B., and Yennie, C. J. (2012) Chemical and biological consequences of oxidatively damaged guanine in DNA. *Free Radical Res.* 46, 420-441.
- (7) Moller, P., Cooke, M. S., Collins, A., Olinski, R., Rozalski, R., and Loft, S. (2012) Harmonising measurements of 8-oxo-7,8-dihydro-2'-deoxyguanosine in cellular DNA and urine. *Free Radical Res.* 46, 541-553.
- (8) Shibutani, S., Takeshita, M., and Grollman, A. P. (1991) Insertion of specific bases during DNA synthesis past the oxidation-damaged base 8-oxodG. *Nature* 349, 431-434.
- (9) Cheng, K. C., Cahill, D. S., Kasai, H., Nishimura, S., and Loeb, L. A. (1992) 8-Hydroxyguanine, an abundant form of oxidative DNA damage, causes G→T and A→C substitutions. *J. Biol. Chem.* 267, 166-172.
- (10) Wood, M. L., Dizdaroglu, M., Gajewski, E., and Essigmann, J. M. (1990) Mechanistic studies of ionizing radiation and oxidative mutagenesis: genetic effects of a single 8-hydroxyguanine (7-hydro-8-oxoguanine) residue inserted at a unique site in a viral genome. *Biochemistry* 29, 7024-7032.

- (11) Klein, J. C., Bleeker, M. J., Saris, C. P., Roelen, H. C., Brugghe, H. F., van den Elst, H., van der Marel, G. A., van Boom, J. H., Westra, J. G., Kriek, E., and et al. (1992) Repair and replication of plasmids with site-specific 8-oxodG and 8-AAFdG residues in normal and repair-deficient human cells. *Nucleic Acids Res.* *20*, 4437-4443.
- (12) David, S. S., O'Shea, V. L., and Kundu, S. (2007) Base-excision repair of oxidative DNA damage. *Nature* *447*, 941-950.
- (13) Steenken, S., Jovanovic, S. V., Bietti, M., and Bernhard, K. (2000) The trap depth (in DNA) of 8-oxo-7,8-dihydro-2'-deoxyguanosine as derived from electron-transfer equilibria in aqueous solution. *J. Am. Chem. Soc.* *122*, 2373-2374.
- (14) Luo, W., Muller, J. G., Rachlin, E. M., and Burrows, C. J. (2000) Characterization of spiroiminodihydantoin as a product of one-electron oxidation of 8-oxo-7,8-dihydroguanosine. *Org. Lett.* *2*, 613-616.
- (15) Luo, W., Muller, J. G., Rachlin, E. M., and Burrows, C. J. (2001) Characterization of hydantoin products from one-electron oxidation of 8-oxo-7,8-dihydroguanosine in a nucleoside model. *Chem. Res. Toxicol.* *14*, 927-938.
- (16) Niles, J. C., Wishnok, J. S., and Tannenbaum, S. R. (2004) Spiroiminodihydantoin and guanidinohydantoin are the dominant products of 8-oxoguanosine oxidation at low fluxes of peroxyxynitrite: mechanistic studies with ^{18}O . *Chem. Res. Toxicol.* *17*, 1510-1519.
- (17) Munk, B. H., Burrows, C. J., and Schlegel, H. B. (2008) An exploration of mechanisms for the transformation of 8-oxoguanine to guanidinohydantoin and spiroiminodihydantoin by density functional theory. *J. Am. Chem. Soc.* *130*, 5245-5256.

- (18) Fleming, A. M., Muller, J. G., Dlouhy, A. C., and Burrows, C. J. (2012) Structural context effects in the oxidation of 8-oxo-7,8-dihydro-2'-deoxyguanosine to hydantoin products: electrostatics, base stacking, and base pairing. *J. Am. Chem. Soc.* *134*, 15091-15102.
- (19) Rokhlenko, Y., Geacintov, N. E., and Shafirovich, V. (2012) Lifetimes and reaction pathways of guanine radical cations and neutral guanine radicals in an oligonucleotide in aqueous solutions. *J. Am. Chem. Soc.* *134*, 4955-4962.
- (20) Cui, L., Ye, W., Prestwich, E. G., Wishnok, J. S., Taghizadeh, K., Dedon, P. C., and Tannenbaum, S. R. (2013) Comparative analysis of four oxidized guanine lesions from reactions of DNA with peroxynitrite, singlet oxygen, and γ -radiation. *Chem. Res. Toxicol.* *26*, 195-202.
- (21) Hailer, M. K., Slade, P. G., Martin, B. D., and Sugden, K. D. (2005) Nei deficient *Escherichia coli* are sensitive to chromate and accumulate the oxidized guanine lesion spiroiminodihydantoin. *Chem. Res. Toxicol.* *18*, 1378-1383.
- (22) Mangerich, A., Knutson, C. G., Parry, N. M., Muthupalani, S., Ye, W., Prestwich, E., Cui, L., McFaline, J. L., Mobley, M., Ge, Z., Taghizadeh, K., Wishnok, J. S., Wogan, G. N., Fox, J. G., Tannenbaum, S. R., and Dedon, P. C. (2012) Infection-induced colitis in mice causes dynamic and tissue-specific changes in stress response and DNA damage leading to colon cancer. *Proc. Natl. Acad. Sci. U. S. A.* *109*, E1820-E1829.
- (23) Leipold, M. D., Muller, J. G., Burrows, C. J., and David, S. S. (2000) Removal of hydantoin products of 8-oxoguanine oxidation by the *Escherichia coli* DNA repair enzyme, FPG. *Biochemistry* *39*, 14984-14992.

- (24) Hazra, T. K., Muller, J. G., Manuel, R. C., Burrows, C. J., Lloyd, R. S., and Mitra, S. (2001) Repair of hydantoins, one electron oxidation product of 8-oxoguanine, by DNA glycosylases of *Escherichia coli*. *Nucleic Acids Res.* 29, 1967-1974.
- (25) Krishnamurthy, N., Muller, J. G., Burrows, C. J., and David, S. S. (2007) Unusual structural features of hydantoin lesions translate into efficient recognition by *Escherichia coli* Fpg. *Biochemistry* 46, 9355-9365.
- (26) Leipold, M. D., Workman, H., Muller, J. G., Burrows, C. J., and David, S. S. (2003) Recognition and removal of oxidized guanines in duplex DNA by the base excision repair enzymes hOGG1, yOGG1, and yOGG2. *Biochemistry* 42, 11373-11381.
- (27) Hailer, M. K., Slade, P. G., Martin, B. D., Rosenquist, T. A., and Sugden, K. D. (2005) Recognition of the oxidized lesions spiroiminodihydantoin and guanidinohydantoin in DNA by the mammalian base excision repair glycosylases NEIL1 and NEIL2. *DNA Repair* 4, 41-50.
- (28) Krishnamurthy, N., Zhao, X., Burrows, C. J., and David, S. S. (2008) Superior removal of hydantoin lesions relative to other oxidized bases by the human DNA glycosylase hNEIL1. *Biochemistry* 47, 7137-7146.
- (29) Yeo, J., Goodman, R. A., Schirle, N. T., David, S. S., and Beal, P. A. (2010) RNA editing changes the lesion specificity for the DNA repair enzyme NEIL1. *Proc. Natl. Acad. Sci. U. S. A.* 107, 20715-20719.
- (30) Zhao, X., Krishnamurthy, N., Burrows, C. J., and David, S. S. (2010) Mutation versus repair: NEIL1 removal of hydantoin lesions in single-stranded, bulge, bubble, and duplex DNA contexts. *Biochemistry* 49, 1658-1666.

- (31) Bandaru, V., Zhao, X., Newton, M. R., Burrows, C. J., and Wallace, S. S. (2007) Human endonuclease VIII-like (NEIL) proteins in the giant DNA Mimivirus. *DNA Repair* 6, 1629-1641.
- (32) McKibbin, P. L., Fleming, A. M., Towheed, M. A., Van Houten, B., Burrows, C. J., and David, S. S. (2013) Repair of hydantoin lesions and their amine adducts in DNA by base and nucleotide excision repair. *J. Am. Chem. Soc.* 135, 13851-13861.
- (33) Duarte, V., Muller, J. G., and Burrows, C. J. (1999) Insertion of dGMP and dAMP during *in vitro* DNA synthesis opposite an oxidized form of 7,8-dihydro-8-oxoguanine. *Nucleic Acids Res.* 27, 496-502.
- (34) Korniyushyna, O., Berges, A. M., Muller, J. G., and Burrows, C. J. (2002) In vitro nucleotide misinsertion opposite the oxidized guanosine lesions spiroiminodihydantoin and guanidinohydantoin and DNA synthesis past the lesions using *Escherichia coli* DNA polymerase I (Klenow Fragment). *Biochemistry* 41, 15304-15314.
- (35) Henderson, P. T., Delaney, J. C., Muller, J. G., Neeley, W. L., Tannenbaum, S. R., Burrows, C. J., and Essigmann, J. M. (2003) The hydantoin lesions formed from oxidation of 7,8-dihydro-8-oxoguanine are potent sources of replication errors *in vivo*. *Biochemistry* 42, 9257-9262.
- (36) Delaney, S., Delaney, J. C., and Essigmann, J. M. (2007) Chemical-biological fingerprinting: probing the properties of DNA lesions formed by peroxynitrite. *Chem. Res. Toxicol.* 20, 1718-1729.
- (37) Delaney, S., Neeley, W. L., Delaney, J. C., and Essigmann, J. M. (2007) The substrate specificity of MutY for hyperoxidized guanine lesions *in vivo*. *Biochemistry* 46, 1448-1455.

- (38) Kamiya, H., and Kasai, H. (1995) Formation of 2-hydroxydeoxyadenosine triphosphate, an oxidatively damaged nucleotide, and its incorporation by DNA polymerases. Steady-state kinetics of the incorporation. *J. Biol. Chem.* 270, 19446-19450.
- (39) Maki, H., and Sekiguchi, M. (1992) MutT protein specifically hydrolyses a potent mutagenic substrate for DNA synthesis. *Nature* 355, 273-275.
- (40) Pavlov, Y. I., Minnick, D. T., Izuta, S., and Kunkel, T. A. (1994) DNA replication fidelity with 8-oxodeoxyguanosine triphosphate. *Biochemistry* 33, 4695-4701.
- (41) Einolf, H. J., and Guengerich, F. P. (2001) Fidelity of nucleotide insertion at 8-oxo-7,8-dihydroguanine by mammalian DNA polymerase δ . Steady-state and pre-steady-state kinetic analysis. *J. Biol. Chem.* 276, 3764-3771.
- (42) Hanes, J. W., Thal, D. M., and Johnson, K. A. (2006) Incorporation and replication of 8-oxo-deoxyguanosine by the human mitochondrial DNA polymerase. *J. Biol. Chem.* 281, 36241-36248.
- (43) Brown, J. A., Duym, W. W., Fowler, J. D., and Suo, Z. (2007) Single-turnover kinetic analysis of the mutagenic potential of 8-oxo-7,8-dihydro-2'-deoxyguanosine during gap-filling synthesis catalyzed by human DNA polymerases λ and β . *J. Mol. Biol.* 367, 1258-1269.
- (44) Jarem, D. A., Wilson, N. R., and Delaney, S. (2009) Structure-dependent DNA damage and repair in a trinucleotide repeat sequence. *Biochemistry* 48, 6655-6663.
- (45) Cantor, C. R., Warshaw, M. M., and Shapiro, H. (1970) Oligonucleotide interactions. 3. Circular dichroism studies of the conformation of deoxyoligonucleotides. *Biopolymers* 9, 1059-1077.

- (46) Cavaluzzi, M. J., and Borer, P. N. (2004) Revised UV extinction coefficients for nucleoside-5'-monophosphates and unpaired DNA and RNA. *Nucleic Acids Res.* 32, e13.
- (47) Ye, Y., Muller, J. G., and Burrows, C. J. (2006) Synthesis and characterization of the oxidized dGTP lesions spiroiminodihydantoin-2'-deoxynucleoside-5'-triphosphate and guanidinohydantoin-2'-deoxynucleoside-5'-triphosphate. *J. Org. Chem.* 71, 2181-2184.
- (48) Karwowski, B., Dupeyrat, F., Bardet, M., Ravanat, J. L., Krajewski, P., and Cadet, J. (2006) Nuclear magnetic resonance studies of the 4*R* and 4*S* diastereomers of spiroiminodihydantoin 2'-deoxyribonucleosides: absolute configuration and conformational features. *Chem. Res. Toxicol.* 19, 1357-1365.
- (49) Fleming, A. M., Orendt, A. M., He, Y., Zhu, J., Dukor, R. K., and Burrows, C. J. (2013) Reconciliation of chemical, enzymatic, spectroscopic and computational data to assign the absolute configuration of the DNA base lesion spiroiminodihydantoin. *J. Am. Chem. Soc.* 135, 18191-18204.
- (50) Johnson, K. A., Simpson, Z. B., and Blom, T. (2009) FitSpace explorer: an algorithm to evaluate multidimensional parameter space in fitting kinetic data. *Anal. Biochem.* 387, 30-41.
- (51) Johnson, K. A., Simpson, Z. B., and Blom, T. (2009) Global kinetic explorer: a new computer program for dynamic simulation and fitting of kinetic data. *Anal. Biochem.* 387, 20-29.
- (52) Johnson, K. A. (2009) Chapter 23 Fitting enzyme kinetic data with KinTek global kinetic explorer. *Methods Enzymol.* 467, 601-626.

- (53) Paul, N., Nashine, V. C., Hoops, G., Zhang, P., Zhou, J., Bergstrom, D. E., and Davisson, V. J. (2003) DNA polymerase template interactions probed by degenerate isosteric nucleobase analogs. *Chem. Biol.* 10, 815-825.
- (54) Eckenroth, B. E., Fleming, A. M., Sweasy, J. B., Burrows, C. J., and Doublet, S. (2014) Crystal structure of DNA polymerase beta with DNA containing the base lesion spiroiminodihydroantoin in a templating position. *Biochemistry* 53, 2075-2077.
- (55) Batra, V. K., Shock, D. D., Beard, W. A., McKenna, C. E., and Wilson, S. H. (2012) Binary complex crystal structure of DNA polymerase β reveals multiple conformations of the templating 8-oxoguanine lesion. *Proc. Natl. Acad. Sci. U. S. A.* 109, 113-118.
- (56) Batra, V. K., Beard, W. A., Hou, E. W., Pedersen, L. C., Prasad, R., and Wilson, S. H. (2010) Mutagenic conformation of 8-oxo-7,8-dihydro-2'-dGTP in the confines of a DNA polymerase active site. *Nat. Struct. Mol. Biol.* 17, 889-890.
- (57) Patra, A., Nagy, L. D., Zhang, Q., Su, Y., Muller, L., Guengerich, F. P., and Egli, M. (2014) Kinetics, structure, and mechanism of 8-oxo-7,8-dihydro-2'-deoxyguanosine bypass by human DNA polymerase η . *J. Biol. Chem.* 289, 16867-16882.
- (58) Lee, H. R., Helquist, S. A., Kool, E. T., and Johnson, K. A. (2008) Importance of hydrogen bonding for efficiency and specificity of the human mitochondrial DNA polymerase. *J. Biol. Chem.* 283, 14402-14410.
- (59) Dahlberg, M. E., and Benkovic, S. J. (1991) Kinetic mechanism of DNA polymerase I (Klenow fragment): identification of a second conformational change and evaluation of the internal equilibrium constant. *Biochemistry* 30, 4835-4843.
- (60) Tsai, Y. C., and Johnson, K. A. (2006) A new paradigm for DNA polymerase specificity. *Biochemistry* 45, 9675-9687.

- (61) Bakhtina, M., Roettger, M. P., Kumar, S., and Tsai, M. D. (2007) A unified kinetic mechanism applicable to multiple DNA polymerases. *Biochemistry* 46, 5463-5472.
- (62) Hanes, J. W., and Johnson, K. A. (2007) A novel mechanism of selectivity against AZT by the human mitochondrial DNA polymerase. *Nucleic Acids Res.* 35, 6973-6983.
- (63) Eger, B. T., and Benkovic, S. J. (1992) The minimal kinetic mechanism for misincorporation by DNA polymerase I (Klenow fragment). *Biochemistry* 31, 9227-9236.
- (64) Neeley, W. L., Delaney, S., Alekseyev, Y. O., Jarosz, D. F., Delaney, J. C., Walker, G. C., and Essigmann, J. M. (2007) DNA polymerase V allows bypass of toxic guanine oxidation products in vivo. *J. Biol. Chem.* 282, 12741-12748.
- (65) Chen, X., Fleming, A. M., Muller, J. G., and Burrows, C. J. (2013) Endonuclease and exonuclease activities on oligodeoxynucleotides containing spiroiminodihydantoin depend on the sequence context and the lesion stereochemistry. *New J. Chem.* 37, 3440-3449.
- (66) Jia, L., Shafirovich, V., Geacintov, N. E., and Broyde, S. (2007) Lesion specificity in the base excision repair enzyme hNei1: modeling and dynamics studies. *Biochemistry* 46, 5305-5314.
- (67) Yanofsky, C., Cox, E. C., and Horn, V. (1966) The unusual mutagenic specificity of an *E. coli* mutator gene. *Proc. Natl. Acad. Sci. U. S. A.* 55, 274-281.
- (68) Tajiri, T., Maki, H., and Sekiguchi, M. (1995) Functional cooperation of MutT, MutM and MutY proteins in preventing mutations caused by spontaneous oxidation of guanine nucleotide in *Escherichia coli*. *Mutat. Res., DNA Repair* 336, 257-267.
- (69) Kamiya, H., Ishiguro, C., and Harashima, H. (2004) Increased A:T→C:G mutations in the mutT strain upon 8-hydroxy-dGTP treatment: direct evidence for MutT involvement in the prevention of mutations by oxidized dGTP. *J. Biochem.* 136, 359-362.

- (70) Hori, M., Suzuki, T., Minakawa, N., Matsuda, A., Harashima, H., and Kamiya, H. (2011) Mutagenicity of secondary oxidation products of 8-oxo-7,8-dihydro-2'-deoxyguanosine 5'-triphosphate (8-hydroxy-2'- deoxyguanosine 5'-triphosphate). *Mutat. Res.* 714, 11-16.
- (71) Loeb, L. A., Essigmann, J. M., Kazazi, F., Zhang, J., Rose, K. D., and Mullins, J. I. (1999) Lethal mutagenesis of HIV with mutagenic nucleoside analogs. *Proc. Natl. Acad. Sci. U. S. A.* 96, 1492-1497.
- (72) Perales, C., Martin, V., and Domingo, E. (2011) Lethal mutagenesis of viruses. *Curr. Opin. Virol.* 1, 419-422.
- (73) Macpherson, P., Barone, F., Maga, G., Mazzei, F., Karran, P., and Bignami, M. (2005) 8-oxoguanine incorporation into DNA repeats *in vitro* and mismatch recognition by MutS α . *Nucleic Acids Res.* 33, 5094-5105.
- (74) Katafuchi, A., and Nohmi, T. (2010) DNA polymerases involved in the incorporation of oxidized nucleotides into DNA: their efficiency and template base preference. *Mutat. Res.* 703, 24-31.

Chapter 3: Unique Length-Dependent Biophysical Properties of Repetitive DNA[†]

[†]Adapted and Modified From:

Huang, J. and Delaney, S. (2016) Unique Length-Dependent Biophysical

Properties of Repetitive DNA, *J. Phys. Chem. B*, accepted, DOI:

10.1021/acs.jpcc.6b00927

3.1 Abstract

Expansion of a trinucleotide repeat (TNR) sequence is the molecular signature of several neurological disorders. Formation of non-canonical structures by the TNR sequence is proposed to contribute to the expansion mechanism. Furthermore, it is known that the propensity for expansion increases with repeat length. In this work we use calorimetry to describe the thermodynamic parameters (ΔH , $T\Delta S$, and ΔG) of the non-canonical stem-loop hairpins formed by the TNR sequences $(CAG)_n$ and $(CTG)_n$, and also the canonical $(CAG)_n/(CTG)_n$ duplexes, for $n = 6-14$. Using a thermodynamic cycle, we calculated the same thermodynamic parameters describing the process of converting from non-canonical stem-loop hairpins to canonical duplex. In addition to these thermodynamic analyses, we used spectroscopic techniques to determine the rate at which the non-canonical structures convert to duplex, and the activation enthalpy ΔH^\ddagger describing this process. We report that the thermodynamic parameters of unfolding the stem-loop $(CTG)_n$ and $(CAG)_n$ hairpins, along with the thermodynamic and kinetic properties of hairpin to duplex conversion, do not proportionally correspond to the increase of length, but rather show a unique pattern that depends on whether the sequence has an even or odd number of repeats.

3.2 Introduction

Trinucleotide repeat (TNR) expansion has been identified as the molecular basis for numerous neurodegenerative disorders including Friedrich's ataxia, Myotonic dystrophy, Spinocerebellar ataxias, Fragile X syndrome, Kennedy's disease, and Huntington's disease (HD).¹⁻

⁴ For example, HD is caused by the expansion of a $(CAG)_n/(CTG)_n$ TNR sequence in exon 1 of the *huntingtin* gene. Healthy individuals have fewer than 35 repeats; 36-39 repeats is a pre-mutation range which can expand further upon transmission to offspring; and the disease state is characterized by 40 or more repeats.⁵

Although the mechanism(s) of TNR expansion are still not fully understood, DNA replication- and DNA repair-dependent mechanisms have been proposed.⁴ In both instances, formation of non-canonical secondary structures, such as stem-loop hairpins, is proposed to contribute to the expansion mechanism. In a replication-dependent expansion, stem-loop hairpin formation is proposed in a polymerase-slippage model^{6,7} or a replication-restart model.^{8,9} In a repair-dependent event, such as repair of the oxidative DNA lesion 8-oxo-7,8-dihydroguanine (8-oxoG),^{10,11} stem-loop hairpins are formed and can be ligated into DNA, which leads to the expansion.¹² In either replication- or repair-dependent expansion, the thermodynamic and kinetic stabilities of the TNR stem-loop hairpins are crucial factors. Expansion is expected when these stem-loop hairpins are stable and persist on a biologically-relevant timescale. In contrast, if the stem-loop hairpins convert to canonical duplex with the complementary DNA, expansion is not expected to occur.

Previous spectroscopic and calorimetric studies of TNR stem-loop hairpins revealed that the thermodynamic stability of these structures depends on both the TNR sequence and the number of repeats.¹³⁻¹⁵ Notably, a common procedure in these prior studies was to examine DNA substrates

in which the number of TNR repeats was incrementally increased by 5 or 10. While this methodology allows for a large range in the number of TNR repeats, these studies lack the resolution to observe differences that might occur with smaller changes in repeat number.

For this reason, we previously examined the stability of a series of (CTG)_n stem-loop hairpins in which the size of the hairpin was varied in increments of *one* repeat (within the range $n = 6-14$).¹⁶ Using spectroscopic techniques and van't Hoff analysis, we observed a pattern whereby the enthalpic stability (ΔH) of the (CTG)_n stem-loop hairpins generally increases with repeat number; however, instead of increasing linearly with the number of repeats, an apparent stepping is observed. Indeed, ΔH is comparable for the pairs of stem-loop hairpins when $n = \text{even}$ and $n = \text{odd}$ (throughout this manuscript “odd” is defined as “even + 1”). Notably, a plateau in ΔH was observed with greater than 10 repeats due to the failure of the melting transition to follow a two-state model.¹⁶ Instead of melting directly from stem-loop hairpin to an unstructured form, the DNA exists as a population of structural intermediates throughout the melting transition. Indeed, a limitation of van't Hoff analysis is the required assumption that the melting transition follows a two-state model. In addition to determining ΔH , our previous structural and biophysical studies showed that when $n = \text{odd}$ one repeat overhangs the end of the stem.¹⁶ The presence of this overhang shortens the lifetime of the non-canonical structure and facilitates conversion to the canonical CAG/CTG duplex.

To circumvent the limitation of the model-dependent van't Hoff analysis used in our previous study, in this work we used differential scanning calorimetry (DSC) to characterize the thermodynamic properties of (CTG)_n stem-loop hairpins ($n = 6-14$). In addition, we also studied the complementary (CAG)_n stem-loop hairpins ($n = 6-14$). DSC allows for direct measurement of the heat supplied to or released from a system during a melting transition and provides a model-

independent measure of the thermodynamic properties of the stem-loop hairpins.^{17,18} Furthermore, in this work we also define the thermodynamic properties of the canonical $(CAG)_n/(CTG)_n$ duplex and the conversion from non-canonical stem-loop hairpins to canonical duplex. In addition to these thermodynamic analyses, we performed kinetic studies to determine the rates at which complementary $(CAG)_n$ and $(CTG)_n$ stem-loop hairpins convert to canonical $(CAG)_n/(CTG)_n$ duplexes, as well as the activation enthalpies (ΔH^\ddagger). We find that the rate of conversion to canonical duplex is significantly faster when n is odd due to a lower ΔH^\ddagger . Taken together, our data provide a thermodynamic and kinetic description of the non-canonical hairpin structures involved in the proposed expansion mechanisms. Furthermore, our data can be related to the clinically observed phenomenon that the propensity for expansion to occur increases with repeat length, and suggest that the TNR expansion caused by formation of non-canonical structures is driven by kinetic factors rather than thermodynamics.

3.3 Experimental Procedure

3.3.1 Oligonucleotide Synthesis and Purification

All oligonucleotides were synthesized using standard phosphoramidite chemistry with a BioAutomation MerMade 4 DNA/RNA synthesizer. Oligonucleotides were purified using a Dynamax Microsorb C18 reverse phase HPLC column (10 × 250 mm) according to published procedures.¹⁹ Oligonucleotide concentrations were determined at 90 °C using the molar extinction coefficient estimated for single-stranded DNA²⁰ using a Beckman Coulter DU800 UV-visible spectrophotometer equipped with a Peltier thermoelectric device. The identity of the oligonucleotides was confirmed by electrospray ionization mass spectrometry.

3.3.2 Differential Scanning Calorimetry (DSC) Analysis

Calorimetric experiments were performed using a TA Instruments Nano DSC III. The (CTG)_n sequences were at concentrations of 50 μM and the (CAG)_n sequences were at concentrations of 100 μM. All oligonucleotides were suspended in 20 mM sodium phosphate, 100 mM NaCl, pH 7.0. Both the oligonucleotide samples and reference buffer sample were degassed *in vacuo* for 15 min at 25 °C before analysis. All data were recorded with TA Instrument Nano DSCRun software version 4.2.6. Data were obtained by continuously monitoring the excess power required to maintain both sample cell and reference cell at the same temperature. The samples were heated from 0 °C to 105 °C followed by cooling from 105 °C to 0 °C, both at 1.0 °C/min. The sample equilibrated for 10 min at 0 and 105 °C between each cooling and heating cycle, respectively. The resulting thermograms display excess heat capacity as a function of temperature. A total of 12 thermograms were obtained for each DNA sequence. A buffer reference was analyzed using the same procedure described above and the thermograms were corrected using this background. Further analyses were performed on TA NanoAnalyze software version 3.5.1 and thermodynamic parameters were obtained from the forward scans. The thermograms were normalized for concentration and a baseline correction was performed using a 5th order polynomial baseline (Figure 3.1). The ΔH was obtained by integrating the area under the thermogram curve. The melting temperature was determined as the T_{\max} of each transition. In cases where there is overlap between the lower and higher temperature transitions, the data were deconvoluted using Origin software to determine ΔH for the transition of interest (Figure 3.2).

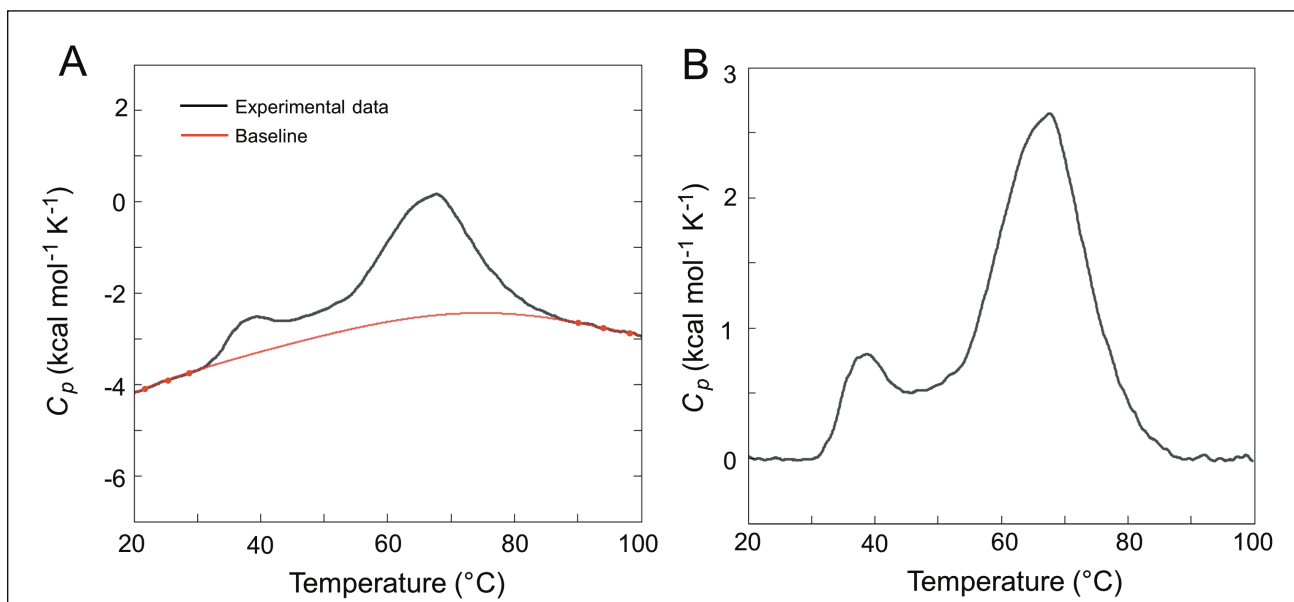
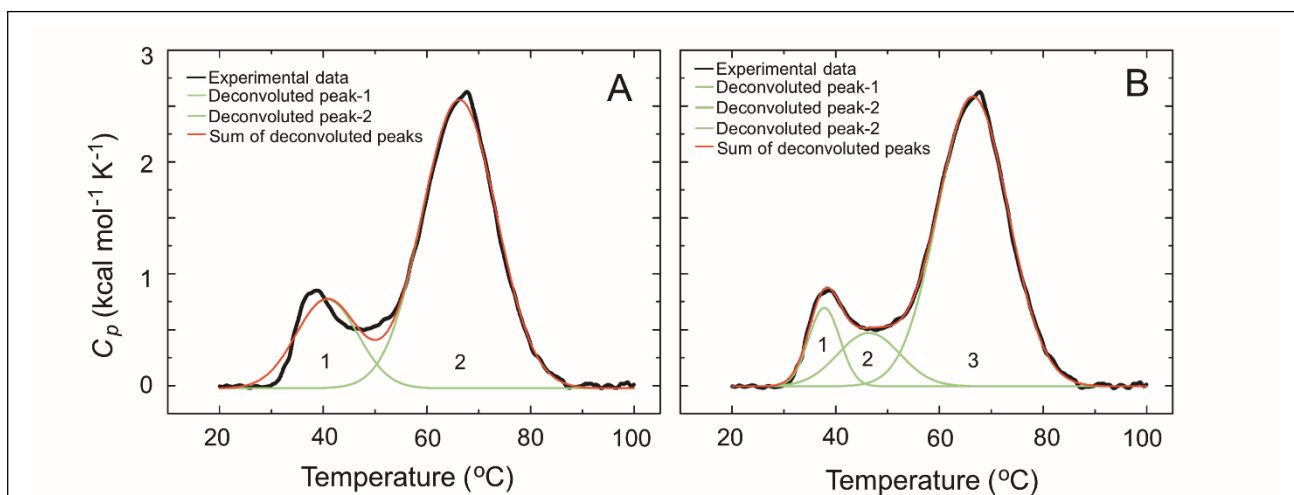


Figure 3.1 Representative example (CTG)₉ of creating 5th order polynomial baseline for overlapped transitions. (A) Three points were placed either prior or post the transitions on the concentration normalized thermogram after buffer scan subtraction. (B) The obtained thermogram after baseline correction.



C

| Deconvolution Method | peak-1 ΔH (kcal/mol) | peak-2 ΔH (kcal/mol) | peak-3 ΔH (kcal/mol) | Total ΔH (kcal/mol) |
|--------------------------|---------------------------------|---------------------------------|---------------------------------|--------------------------------|
| 2-peak deconvolution (A) | 11.9 | 46.2 | / | 58.1 |
| 3-peak deconvolution (B) | 5.3 | 7.2 | 44.2 | 56.7 |

Figure 3.2 Example for (CTG)₉ of deconvolution of overlapped transitions. (A) Experimental data was deconvoluted into two peaks. (B) Experimental data was deconvoluted into 3 peaks. (C) Summary of ΔH using either 2 or 3 peaks in deconvolution. Analysis was carried out assuming the low transition is homoduplex to hairpin transition described in Results and Discussion.

A flash-cooling experiment was also performed for (CTG)₉. The sample was heated to 90 °C for 5 min and immediately placed in an ice/water bath for at least 30 min before loading into the DSC sample cell, which was pre-equilibrated at 4 °C.

3.3.3 Native Polyacrylamide Gel Electrophoresis

Following analysis by DSC, the samples were retrieved and 15 µL of each sample was added to 30,000 cpm of the corresponding dried radiolabeled oligonucleotide. In all cases, the amount of radiolabeled DNA is less than 5 pmol and the change in concentration is negligible. The samples were heated at 105 °C for 15 min and then the temperature was decreased at a rate of 1 °C/min to the final temperature of 10 °C. Fifteen µL of native loading dye (15 % Ficoll, 0.25 % bromophenol blue and 0.25 % xylene cyanol) were added and 3,000 cpm of each sample was resolved by 12 % native polyacrylamide gel electrophoresis (PAGE), where the gel was pre-run at 100 V at 4 °C for at least 1 h, and the gel was run at 35 V at 4 °C for 12 h and visualized by Phosphorimagery.

3.3.4 Kinetics of Hairpin to Duplex Conversion

The (CTG)_n and (CAG)_n oligonucleotides (2 µM each) were separately incubated at 37 °C for 20 min followed by mixing of 162 µL of each oligonucleotide. Using a Beckman Coulter DU800 UV-visible spectrophotometer the absorbance at 260 nm was monitored as a function of time while the sample temperature was maintained at 37 °C. Data were collected until the absorbance remained constant for at least 10 min. The absorbance values were normalized such that the Y-axis represents the fraction of hairpin remaining.²¹ The data were fit with KaleidaGraph software to a second-order equation²² that describes the rate of duplex formation from two

equimolar complementary hairpins with a 15 s time correction²³ to account for the time lag between mixing the two oligonucleotides and beginning to monitor the absorbance.

$$I = I_0 \frac{\epsilon_{duplex}}{\epsilon_{CTG} + \epsilon_{CAG}} + \left(1 - \frac{\epsilon_{duplex}}{\epsilon_{CTG} + \epsilon_{CAG}}\right) \left(\frac{I_0}{1 + [I_0 / (\epsilon_{CTG} + \epsilon_{CAG})]k(t-15)}\right) \quad (1)$$

where I is the absorbance at any time point, I_0 is the absorbance at the start of the reaction, ϵ_{duplex} , ϵ_{CTG} and ϵ_{CAG} are the ϵ_{260} for the CAG/CTG duplex, CTG and CAG hairpins, respectively, k is the rate constant, and t is the reaction time in seconds. At least 5 replicate experiments were performed for each hairpin to duplex conversion and a representative graph of hairpin fraction remaining versus time is provided.

The activation enthalpy (ΔH^\ddagger) describing hairpin to duplex conversion can be determined by performing the same experiment described above at multiple temperatures, where all of the temperatures are below the melting temperature of the hairpin. For these experiments the (CTG)_n and (CAG)_n oligonucleotides were diluted to the same concentration such that after mixing of the complementary sequences, the final UV-visible absorbance was ~0.3-0.4 (~1-2 μ M). The resulting data were fit using the Eyring equation to determine ΔH^\ddagger .^{24,25}

$$\ln \frac{k}{T} = -\frac{\Delta H^\ddagger}{R} \cdot \frac{1}{T} + \ln \frac{k_B}{h} + \frac{\Delta S^\ddagger}{R} \quad (2)$$

where k is the rate constant, T is the absolute temperature, R is the gas constant, k_B is the Boltzmann constant, h is Planck's constant, and ΔH^\ddagger and ΔS^\ddagger are the activation enthalpy and activation entropy, respectively.

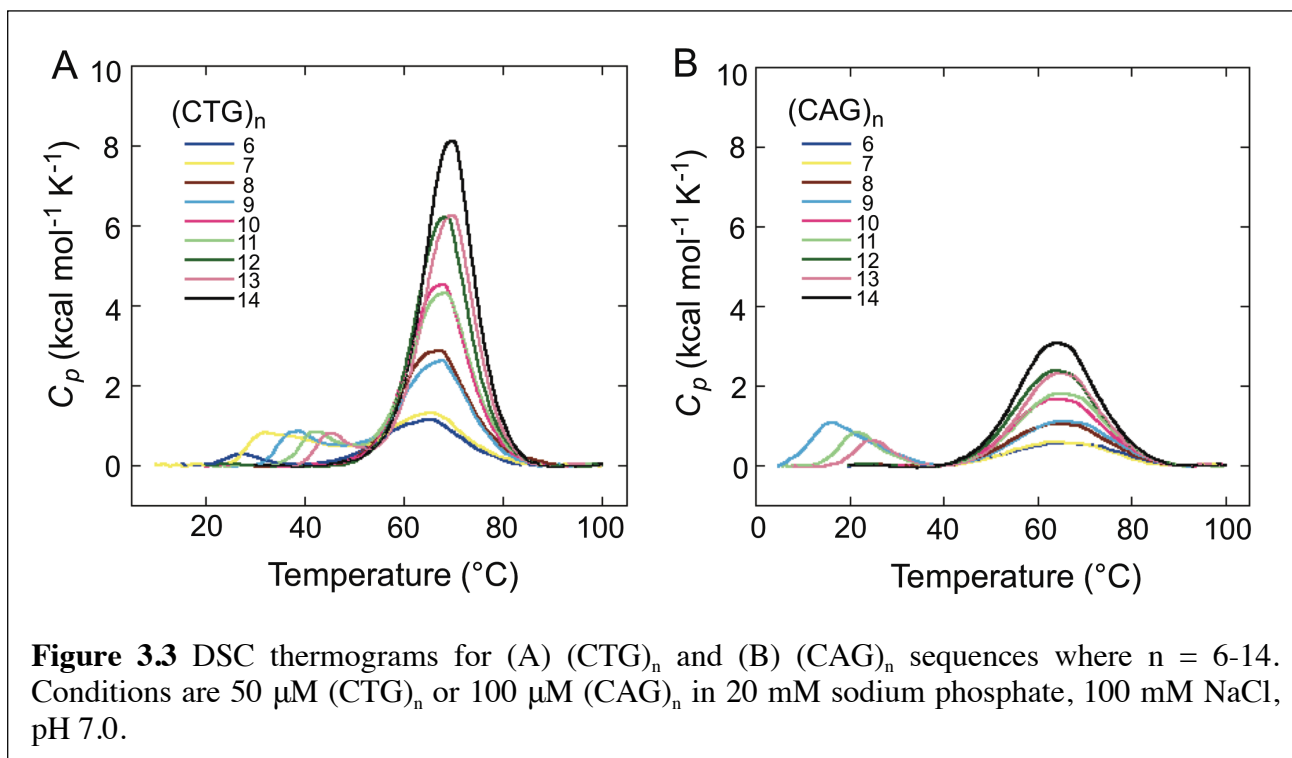
3.4 Results and Discussion

In this work we used a series of (CAG)_n and (CTG)_n DNA oligonucleotides where n = 6-14 to provide a comprehensive calorimetric and thermodynamic analysis of: (1) the stability of

non-canonical stem-loop hairpins formed by the single-stranded oligonucleotides, (2) the stability of canonical $(CAG)_n/(CTG)_n$ duplexes, and (3) the factors affecting the conversion of the non-canonical structures to canonical duplex. In addition, we performed a spectroscopy-based kinetic analysis of the rates at which the non-canonical stem-loop hairpins convert to canonical duplexes and report ΔH^\ddagger for this process.

3.4.1 DSC Characterization of $(CAG)_n$ and $(CTG)_n$ Stem-Loop Hairpins

Using DSC we obtained thermograms characterizing the melting of the $(CTG)_n$ and $(CAG)_n$ samples under physiologically relevant salt conditions (Figure 3.3). In the thermograms excess heat capacity is plotted as a function of temperature. As seen for the $(CTG)_n$ sequences (Figure 3.3A), when $n = 6-14$, a transition centered at ~ 65 °C is observed in each thermogram; furthermore, when $n = 6, 7, 8, 9, 11,$ or 13 , a second transition is observed and is centered at $25-45$ °C. In contrast, when $n = 10, 12,$ or 14 this lower temperature transition is not observed. Similar behavior is observed for the $(CAG)_n$ sequence when $n = 6-14$ (Figure 3.3B). When $n = 6-14$, all thermograms for $(CAG)_n$ display a transition centered at ~ 65 °C, and only when $n = 9, 11,$ or 13 , a lower temperature transition is observed and is centered at $15-25$ °C. Although previous studies only reported single transition for $(CTG)_n$ or $(CAG)_n$ hairpins of similar length,¹⁵ these work did not include repeat number of $n = 7$ and 9 , where the lower temperature transitions are most prevalent in our study.



3.4.2 Electrophoretic Mobility of $(CAG)_n$ and $(CTG)_n$ by Native PAGE

To determine whether the two transitions observed by DSC are due to the melting of two different structures, we performed native PAGE analysis for both $(CAG)_n$ and $(CTG)_n$ (Figure 3.4). For $(CTG)_n$ two distinct species (labeled species A and B) are observed when n is odd and the relative intensity of the two species varies depending on n (Figure 3.4A). While species B is prevalent when n is odd, it is much less prevalent when n is even. Similar results were reported previously for $(CTG)_n$ when $n = 2-10$.²⁶ For $(CAG)_n$ we observe only one species by native PAGE, except $n = 13$ when two species are observed (Figure 3.4B).

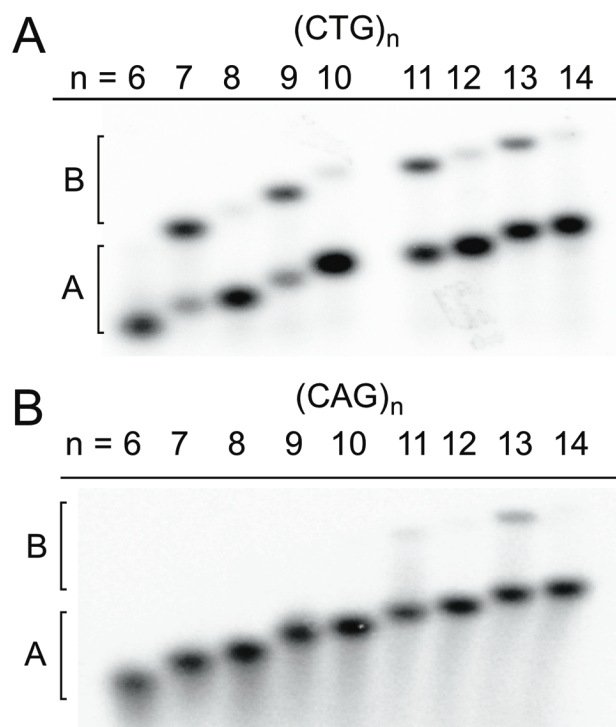
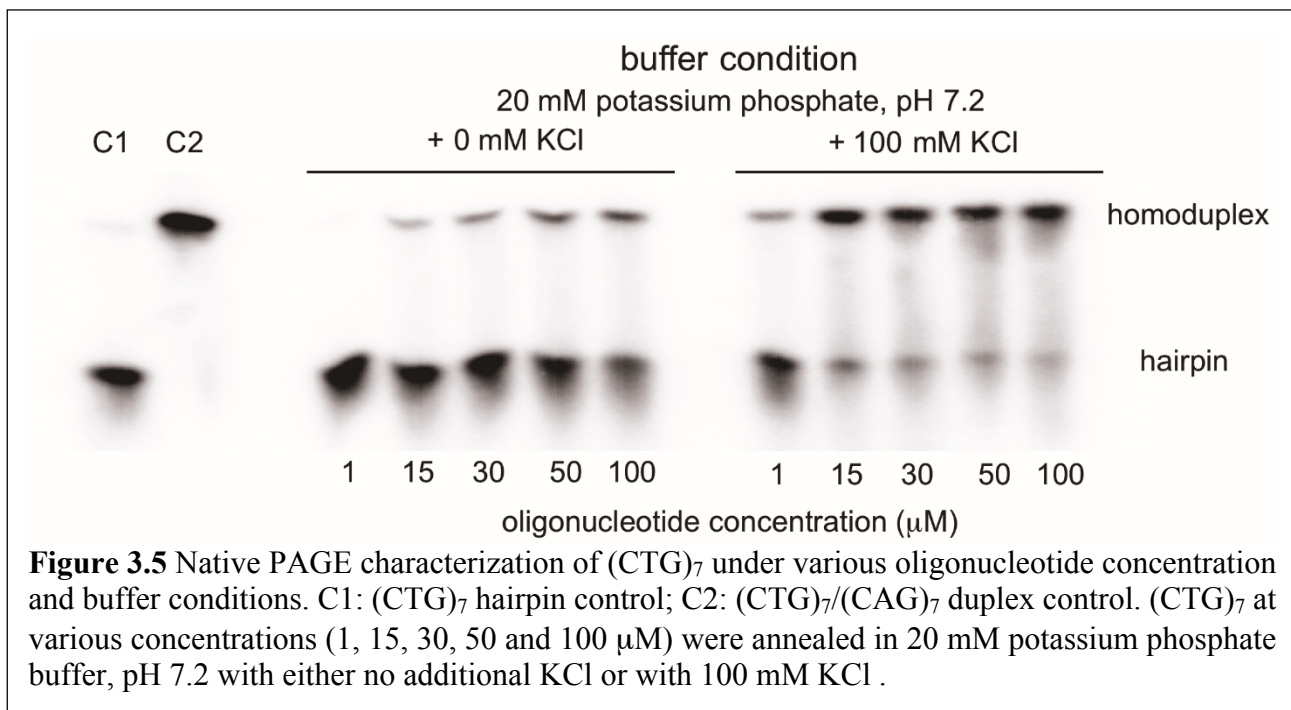


Figure 3.4 Native PAGE gel for (A) (CTG)_n sequences (n = 6-14) at 50 μM; (B) (CAG)_n sequences (n = 6-14) at 100 μM. In all cases DNA was in 20 mM sodium phosphate, 100 mM NaCl, pH 7.0.

Using (CTG)₇ as a representative sample, the electrophoretic mobility of species A and B were compared to authentic standards. Native PAGE reveals that species A of (CTG)₇ co-migrates with a (CTG)₇ stem-loop hairpin control while species B co-migrates with a (CTG)₇/(CAG)₇ heteroduplex control (Figure 3.5). As described previously by Lam, we assign species B as (CTG)₇/(CTG)₇ homoduplex due to its slower electrophoretic mobility relative to (CTG)₇ stem-loop hairpin.²⁶ Further analysis showed that the amount of homoduplex increases as a function of both oligonucleotide and salt concentration (Figure 3.5). It is noteworthy that using a flash-cooling annealing procedure prior to native PAGE or DSC, a process which is expected to trap any kinetically favored species, did not influence the amount of stem-loop hairpin or homoduplex (data not shown). Consistent with this result, previous studies using the dimerization initiation site of

HIV-1 genomic RNA showed that concentration of the RNA, rather than the annealing procedure, is the major parameter dictating stem-loop hairpin formation.²⁷



Previous work from our¹⁶ and other laboratories^{13,15,21} has shown that (CTG)_n and (CAG)_n stem-loop hairpins show very little change in melting temperature (T_m) as a function of length. A lack of dependence on length is also seen for the higher temperature transitions in DSC; this observation suggests that the transitions centered at ~65 °C are due to the melting of the stem-loop hairpins. Thus, the length-dependent transitions at lower temperature are ascribed to melting of homoduplexes. Shown in Table 3.1 are the T_m values obtained by DSC for the homoduplexes and stem-loop hairpins. For the (CTG)_n sequences the T_m values of the stem-loop hairpins vary from 62.8-69.2 °C. The T_m values for the homoduplexes increase with n over a range of ~26-45 °C. For the (CAG)_n sequences there is very little variation in the T_m of the stem-loop hairpins when n = 6-14; the average T_m is 64.3 °C. The T_m for (CAG)_n homoduplexes increase as a function of length over a range of ~16-25 °C.

Table 3.1 DSC-Derived Melting Temperatures and Thermodynamic Parameters for (CTG)_n and (CAG)_n Stem-Loop Hairpins.^{1,2}

| Sequence | T _m Hairpin (°C) | T _m Homoduplex (°C) | ΔH (kcal mol ⁻¹) | TΔS ³ (kcal mol ⁻¹) | ΔG ³ (kcal mol ⁻¹) |
|---------------------|-----------------------------|--------------------------------|------------------------------|--|---|
| (CTG) ₆ | 65.4 | 27.8 | 22.4 ± 0.7 | 20.4 ± 0.6 | 1.9 ± 0.2 |
| (CTG) ₇ | 66.1 | 32.0 | 24.9 ± 2.3 | 22.8 ± 2.1 | 2.1 ± 0.2 |
| (CTG) ₈ | 67.3 | - | 48.1 ± 1.4 | 43.7 ± 1.2 | 4.3 ± 0.1 |
| (CTG) ₉ | 67.7 | 38.7 | 45.0 ± 1.0 | 40.9 ± 0.9 | 4.1 ± 0.1 |
| (CTG) ₁₀ | 68.2 | - | 67.7 ± 1.8 | 61.4 ± 1.6 | 6.2 ± 0.1 |
| (CTG) ₁₁ | 68.5 | 42.0 | 64.7 ± 1.2 | 58.6 ± 0.9 | 6.0 ± 0.1 |
| (CTG) ₁₂ | 68.8 | - | 86.4 ± 2.5 | 78.4 ± 2.2 | 8.0 ± 0.2 |
| (CTG) ₁₃ | 70.1 | 45.1 | 82.1 ± 1.0 | 74.1 ± 0.9 | 7.9 ± 0.1 |
| (CTG) ₁₄ | 70.0 | - | 108 ± 1 | 97.3 ± 0.6 | 10.4 ± 0.1 |
| (CAG) ₆ | 64.3 | - | 14.2 ± 0.6 | 13.1 ± 0.5 | 1.1 ± 0.1 |
| (CAG) ₇ | 64.1 | - | 15.1 ± 1.1 | 13.9 ± 1.0 | 1.2 ± 0.1 |
| (CAG) ₈ | 63.9 | - | 25.6 ± 0.6 | 23.5 ± 0.6 | 2.1 ± 0.1 |
| (CAG) ₉ | 65.2 | 16.4 | 25.0 ± 2.7 | 22.9 ± 2.4 | 2.1 ± 0.3 |
| (CAG) ₁₀ | 64.1 | - | 37.4 ± 1.1 | 34.4 ± 0.9 | 3.0 ± 0.1 |
| (CAG) ₁₁ | 64.3 | 21.6 | 38.6 ± 3.0 | 35.3 ± 2.8 | 3.2 ± 0.3 |
| (CAG) ₁₂ | 64.0 | - | 50.1 ± 1.3 | 45.9 ± 1.2 | 4.0 ± 0.1 |
| (CAG) ₁₃ | 65.0 | 24.5 | 49.8 ± 1.5 | 45.6 ± 1.2 | 4.2 ± 0.1 |
| (CAG) ₁₄ | 64.1 | - | 62.6 ± 1.4 | 57.7 ± 1.2 | 5.1 ± 0.1 |

¹ In 20 mM sodium phosphate, 100 mM NaCl, pH 7.0.

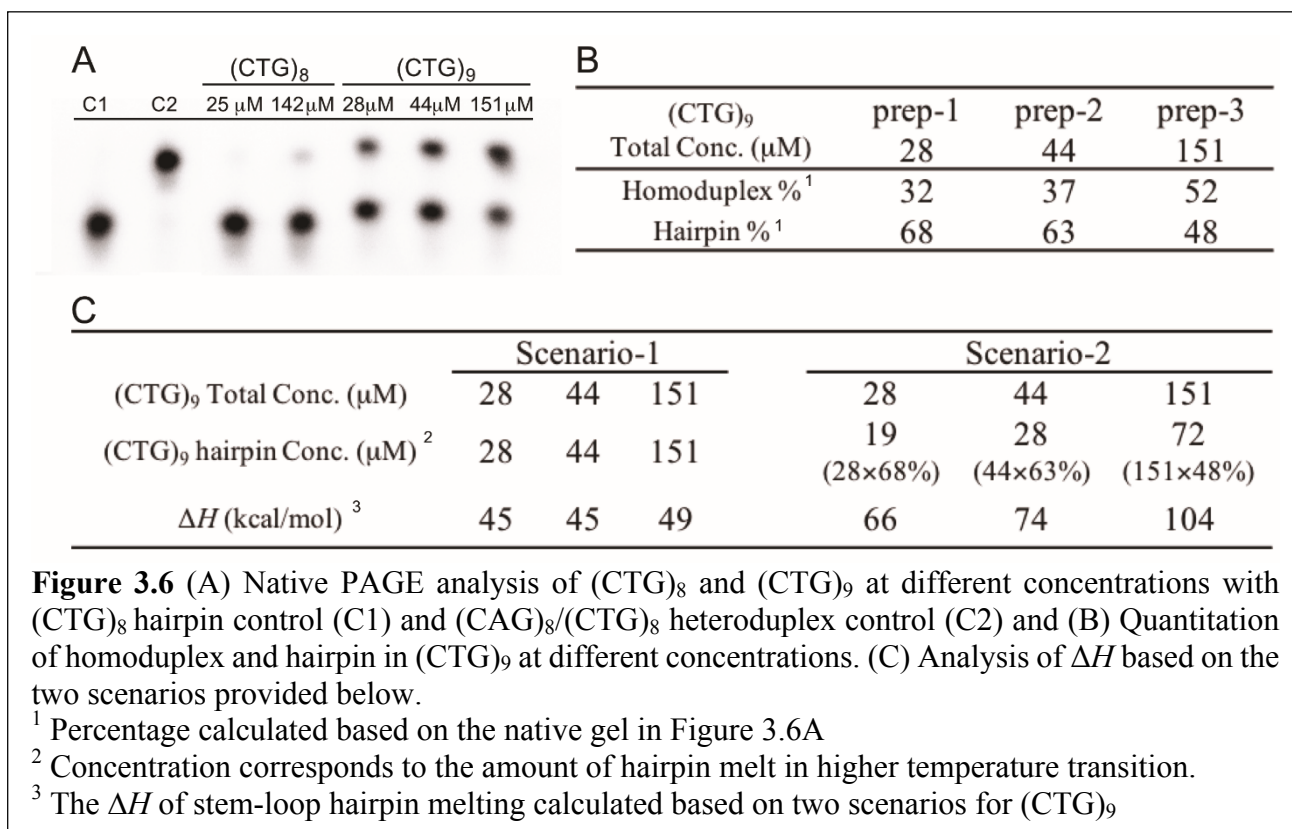
² Errors represent standard deviation from the analysis of three scans of a single sample preparation.

³ Values at 37 °C.

3.4.3 Determining the Nature of Homoduplex Melting

Here we are particularly interested in the thermodynamic parameters describing melting of the stem-loop hairpins, which are the transitions centered at ~65 °C. While the thermodynamic parameters of the homoduplex are not our focus *per se*, the nature of homoduplex melting could affect our analysis of the transition centered at ~65 °C. In order to determine thermodynamic parameters from a DSC thermogram, the concentration of the species undergoing the melting transition must be known. We envision two possible scenarios to describe homoduplex melting: (1) the homoduplex melts and each strand forms a stem-loop hairpin; the stem-loop hairpins subsequently melt to unstructured single-stranded DNA at the transition centered at ~65 °C, or (2)

the homoduplex melts directly to unstructured single-stranded DNA. If scenario 1 is followed, the transition at ~65 °C would correspond to melting of the entire or majority population of DNA in the sample (depending on the extent of overlapping between the two transitions). In contrast, if scenario 2 is followed and the homoduplex melts directly to unstructured single-stranded DNA, the transition at ~65 °C would not involve the entire population but rather the hairpin fraction of the entire DNA population as determined by native gel electrophoresis.



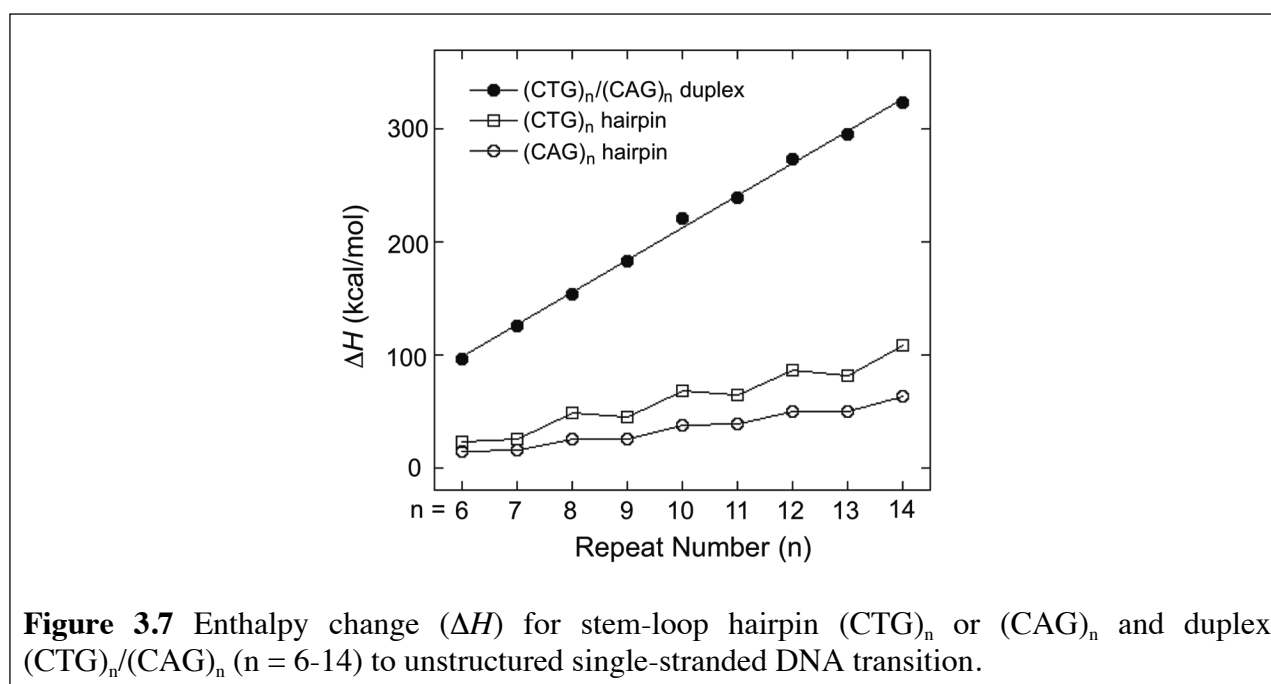
To determine the nature of homoduplex melting we performed DSC and native PAGE experiments as a function of concentration for (CTG)₉ (Figure 3.6). Native PAGE was used to obtain the percentage of stem-loop hairpin and homoduplex present at each concentration (Figure 3.6B). Since both the percentages of stem-loop hairpin and homoduplex are concentration-dependent, based on the amount of stem-loop hairpin and homoduplex, we obtained values for ΔH

corresponding to stem-loop hairpin melting using the DSC thermograms and assuming either scenario 1 (Figure 3.6C) or scenario 2 (Figure 3.6D). It is known that ΔH for stem-loop hairpin melting would be independent of concentration,²⁸ which is the data we obtain using scenario 1 is followed; thus, we conclude that scenario 1 is followed and the homoduplex melts and each strand forms a stem-loop hairpin. Therefore, the remaining DSC thermograms were analyzed assuming that the transition at ~ 65 °C corresponds to melting of a stem-loop hairpin for the entire population of DNA.

3.4.4 Analysis of DSC Thermograms for $(CAG)_n$ and $(CTG)_n$ Stem-Loop Hairpins

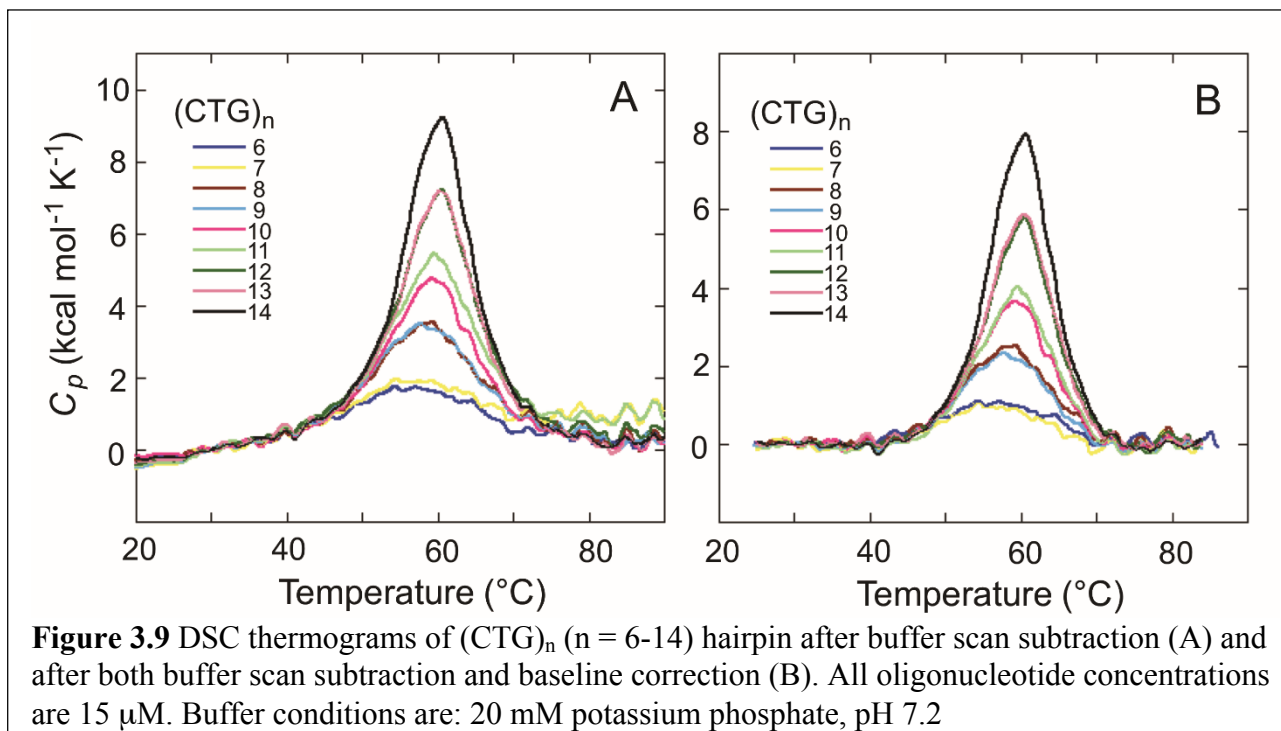
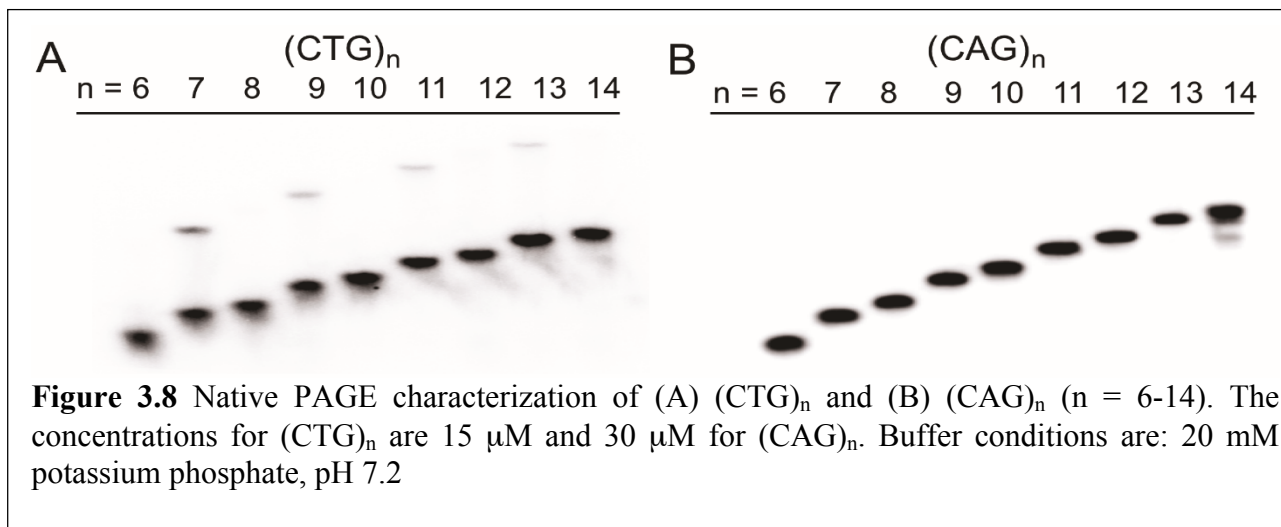
The assignment of the two thermal transitions in DSC allows us to further analyze the thermodynamic parameters for unfolding of $(CAG)_n$ and $(CTG)_n$ stem-loop hairpins. It is straightforward to analyze thermodynamic parameters for samples with only a single transition or with two well-separated transitions. For samples with overlapped transitions in the thermograms, it is best to deconvolute the experimental data into three transitions (Figure 3.2). Notably, ΔH for the hairpin to unstructured transition remains the same regardless of the number of transitions used during the deconvolution (Figure 3.2). Based on the deconvolution results, it is clear that for samples with overlapped transitions, all of the homoduplex has been converted to hairpin before the T_m of the hairpin itself, so the higher temperature transition contains the entire concentration of DNA and simplifies the analysis of thermodynamic parameters. The thermodynamic parameters describing the transition from stem-loop hairpin to unstructured single strand are summarized in Table 3.1. Notably, the reported values describe the process of melting from stem-loop hairpin to unstructured single strand and ΔH , $T\Delta S$, and ΔG are all positive values. We observe that when n is the same, the $(CTG)_n$ stem-loop hairpins are enthalpically (ΔH) stabilized and entropically ($T\Delta S$)

destabilized relative to the $(CAG)_n$ stem-loop hairpins, as the values for ΔH and $T\Delta S$ are greater for the $(CTG)_n$ stem-loop hairpins. It can also be seen that $(CTG)_n$ stem-loop hairpins with $n =$ even have similar thermodynamic parameters, as when $n =$ odd [i.e., $(CTG)_6$ and $(CTG)_7$ have similar thermodynamic parameters, as do $(CTG)_8$ and $(CTG)_9$, $(CTG)_{10}$ and $(CTG)_{11}$, and $(CTG)_{12}$ and $(CTG)_{13}$] (Figure 3.7). The same pattern is observed for the $(CAG)_n$ stem-loop hairpins (Figure 3). Finally, while $n =$ even and $n =$ odd hairpins have similar thermodynamic parameters, there is an overall increase in ΔH , ΔS , and ΔG as a function of repeat length across the series $n = 6-14$.



In order to confirm the validity of our analysis of the thermograms containing overlapping transitions, we also looked for experimental conditions where no homoduplex was observed. It is known that at low concentrations, $(CAG)_n$ and $(CTG)_n$ will form intramolecular stem-loop hairpins exclusively, as confirmed by the hairpin control in our native gel (Figure 3.3). We also found that the homoduplex population was decreased in low salt buffer conditions. Therefore, we repeated the calorimetry experiments for $(CAG)_n$ and $(CTG)_n$ under conditions where no significant amount of homoduplex is detected by native gel electrophoresis (20 mM potassium phosphate, pH 7.2)

(Figure 3.8). As expected, we only observed a single transition for all samples, while the even/odd pattern remains as indicated by the almost superimposed thermograms for the even and odd number of hairpins (Figure 3.9, 3.10). Due to the reduced salt concentration, both the T_m and ΔH values derived from the stem-loop hairpin to unstructured transitions decrease compared to previous analyses (Table 3.2).



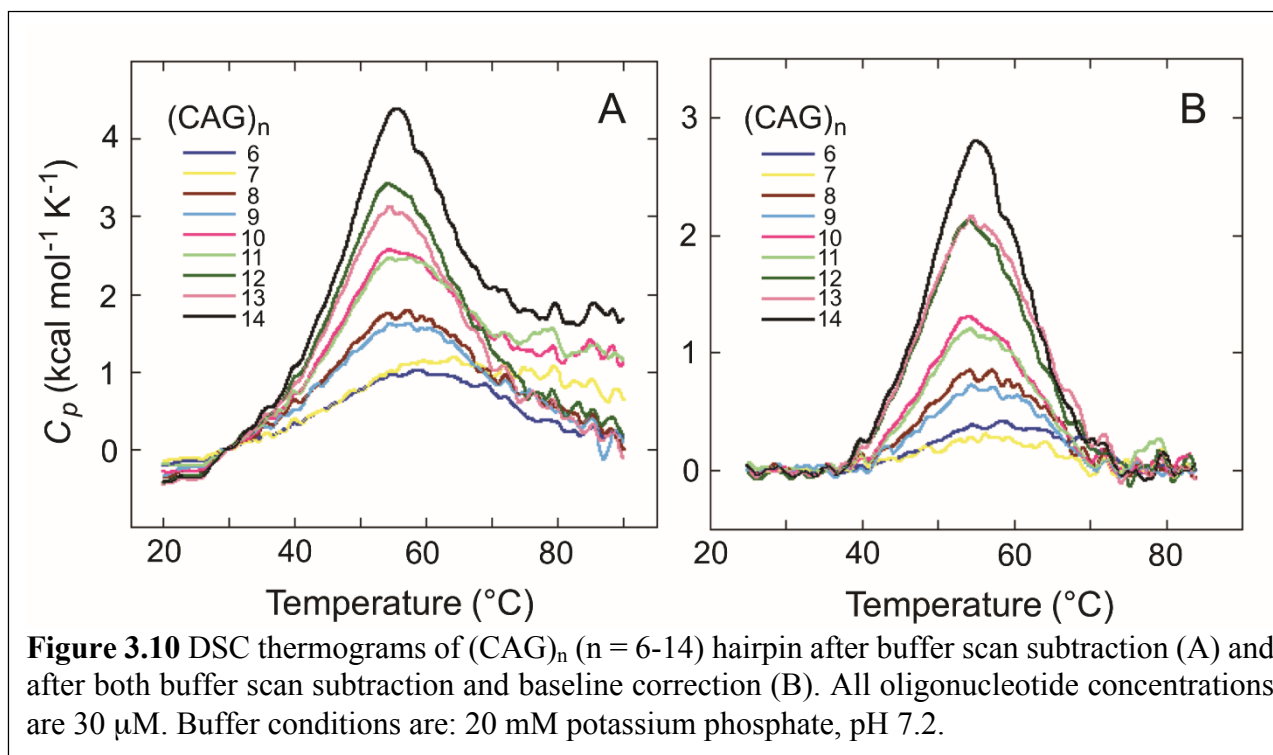


Table 3.2 DSC-Derived Melting Temperatures and Thermodynamic Parameters for $(CTG)_n$ and $(CAG)_n$ Stem-Loop Hairpins under low salt conditions.^{1, 2}

| Sequence | T_m Hairpin ($^{\circ}C$) | ΔH ($kcal\ mol^{-1}$) | $T \Delta S^3$ ($kcal\ mol^{-1}$) | ΔG^3 ($kcal\ mol^{-1}$) |
|--------------|-------------------------------|---------------------------------|-------------------------------------|-----------------------------------|
| $(CTG)_6$ | 56.1 | 18.5 ± 0.8 | 17.5 ± 0.7 | 1.1 ± 0.1 |
| $(CTG)_7$ | 53.9 | 20.0 ± 1.2 | 19.0 ± 1.1 | 1.0 ± 0.1 |
| $(CTG)_8$ | 58.7 | 35.1 ± 1.9 | 32.8 ± 1.8 | 2.3 ± 0.1 |
| $(CTG)_9$ | 58.3 | 37.0 ± 3.3 | 34.6 ± 3.0 | 2.4 ± 0.3 |
| $(CTG)_{10}$ | 59.2 | 50.6 ± 3.3 | 47.2 ± 3.1 | 3.4 ± 0.2 |
| $(CTG)_{11}$ | 59.8 | 52.1 ± 3.0 | 48.5 ± 2.8 | 3.4 ± 0.2 |
| $(CTG)_{12}$ | 60.6 | 68.6 ± 2.3 | 63.7 ± 2.1 | 4.9 ± 0.2 |
| $(CTG)_{13}$ | 60.4 | 69.1 ± 4.3 | 64.2 ± 3.9 | 4.9 ± 0.3 |
| $(CTG)_{14}$ | 60.6 | 85.3 ± 1.7 | 79.3 ± 1.5 | 6.1 ± 0.2 |
| $(CAG)_6$ | 59.0 | 7.6 ± 1.0 | 7.1 ± 0.9 | 0.5 ± 0.1 |
| $(CAG)_7$ | 56.6 | 5.6 ± 0.3 | 5.2 ± 0.4 | 0.3 ± 0.1 |
| $(CAG)_8$ | 55.7 | 15.2 ± 0.4 | 14.3 ± 0.5 | 0.9 ± 0.1 |
| $(CAG)_9$ | 55.3 | 12.6 ± 1.0 | 11.9 ± 1.0 | 0.7 ± 0.1 |
| $(CAG)_{10}$ | 54.2 | 22.3 ± 0.4 | 21.1 ± 0.3 | 1.2 ± 0.1 |
| $(CAG)_{11}$ | 54.2 | 21.5 ± 0.2 | 20.4 ± 0.1 | 1.1 ± 0.1 |
| $(CAG)_{12}$ | 54.1 | 34.1 ± 1.3 | 32.4 ± 1.2 | 1.8 ± 0.1 |
| $(CAG)_{13}$ | 54.3 | 36.8 ± 1.5 | 34.9 ± 1.3 | 2.0 ± 0.1 |
| $(CAG)_{14}$ | 54.8 | 42.6 ± 1.5 | 40.3 ± 1.3 | 2.3 ± 0.2 |

¹ In 20 mM potassium phosphate, pH 7.2.

² Errors represent standard deviation from the analysis of three scans of a single sample preparation.

³ Values at 37 $^{\circ}C$.

Both sets of the calorimetry results suggest that the change in ΔH for hairpin melting is not only proportional to the number of CAG or CTG repeats, but also depends on whether the number is even or odd. Although the low salt conditions used in this experiment are less biologically relevant compared to the experiments described above, the results provide support for our analysis of thermograms with overlapping transitions.

Our results show that $(CTG)_n$ stem-loop hairpins are enthalpically and thermodynamically more stable than the $(CAG)_n$ stem-loop hairpins with the same number of repeats. The notable difference between these stem-loop hairpins is the presence of T•T or A•A mismatches in the stem. Mitas has shown that the T•T mismatches in $(CTG)_n$ stem-loop hairpins are well stacked and can form two hydrogen bonds, while the A•A mismatches in $(CAG)_n$ are less well stacked, and no hydrogen bonds are formed in the mismatch.²⁹ Furthermore, Arnold and co-workers showed by ¹H-NMR that in canonical duplex, the adenines in an A•A mismatch are tilted and pushed apart to avoid overlap of the exocyclic amino groups.³⁰ In contrast, the T•T mismatch results in little distortion of the bases or sugar-phosphate backbone. Both of these observations are consistent with the increased stability of $(CTG)_n$ stem-loop hairpins relative to the corresponding $(CAG)_n$ hairpins.

3.4.5 Thermodynamics of Stem-Loop Hairpin to Duplex Conversion

In addition to the thermodynamic characterization of the $(CAG)_n$ and $(CTG)_n$ stem-loop hairpins, we are interested in the thermodynamic parameters describing conversion of the non-canonical stem-loop hairpins to canonical duplex when the two complementary hairpins are combined. However, the DSC experiments do not allow us to monitor the process directly; a significant amount of the stem-loop hairpins convert to duplex during the dead time of the experiment. Instead, thermodynamic parameters for the stem-loop hairpin to duplex conversion

process were calculated using a thermodynamic cycle, a procedure which was validated previously by Breslauer and co-workers for Ω -DNA structures.³¹ In order to utilize this cycle, the thermodynamic parameters for melting of $(CAG)_n/(CTG)_n$ duplexes are required, and these values were determined by DSC. The thermograms for $(CAG)_n/(CTG)_n$ duplexes all show a single transition where the T_m and ΔH increase as a function of length (Figure 3.7) as compared to the stepwise increase of $(CTG)_n$ and $(CAG)_n$ hairpins. Indeed, a linear increase with repeat length is observed for all the thermodynamic parameters (ΔH , $T\Delta S$, and ΔG) as expected for duplexes of increasing length (Figure 3.11, Table 3.3).

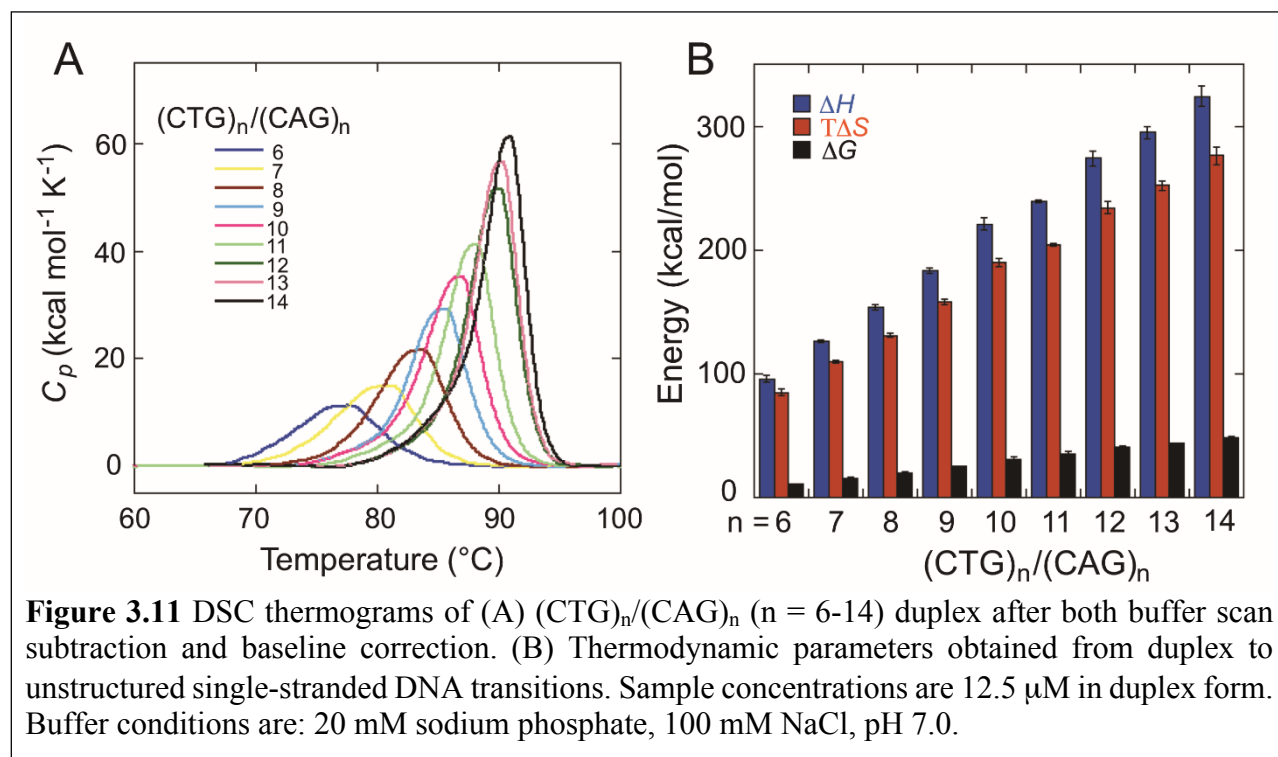


Table 3.3 DSC-Derived Melting Temperatures and Thermodynamic Parameters for (CAG)_n/(CTG)_n Duplexes.^{1,2}

| n | T _m (°C) | ΔH (kcal mol ⁻¹) | T ΔS ³ (kcal mol ⁻¹) | ΔG ³ (kcal mol ⁻¹) |
|----|---------------------|------------------------------|---|---|
| 6 | 77.8 | 96.0 ± 3.1 | 84.7 ± 2.7 | 11.2 ± 0.3 |
| 7 | 81.1 | 126 ± 1 | 110 ± 1 | 15.8 ± 0.2 |
| 8 | 83.8 | 154 ± 2 | 131 ± 2 | 20.3 ± 0.2 |
| 9 | 85.6 | 183 ± 2 | 158 ± 2 | 24.9 ± 0.4 |
| 10 | 86.7 | 221 ± 5 | 190 ± 3 | 31.0 ± 1.5 |
| 11 | 88.2 | 239 ± 1 | 204 ± 1 | 35.1 ± 2.2 |
| 12 | 90.1 | 274 ± 6 | 234 ± 5 | 40.2 ± 1 |
| 13 | 90.3 | 295 ± 5 | 252 ± 4 | 43.4 ± 0.7 |
| 14 | 91.0 | 324 ± 8 | 276 ± 7 | 48.2 ± 1.2 |

¹ In 20 mM sodium phosphate, 100 mM NaCl, pH 7.0.

² Errors represent standard deviation from the analysis of three scans of a single sample preparation.

³ Values at 37 °C.

A thermodynamic cycle was used to calculate the thermodynamic parameters describing the stem-loop hairpin to duplex conversion (Table 3.4, Figure 3.12). It is noteworthy that the values for ΔH, TΔS, and ΔG are all negative, indicating the process is enthalpically and thermodynamically favored but entropically disfavored. Importantly, both ΔH and TΔS display an oscillating pattern across the range of sequence lengths; when n = odd the values are more negative than when n = even. This observation can be rationalized by structural differences in the stem-loop hairpins. When n = odd, one repeat overhangs the end of the stem, and this overhang provides three additional base pairs and additional base stacking when duplex is formed. However, this enthalpic stabilization is not expected for stem-loop hairpins when n = even; in the latter case, the stem has a blunt end with no overhang. Nevertheless, for all values of n, the ΔH term is always more negative than TΔS. Thus, the process of stem-loop hairpin to duplex conversion is enthalpically driven. Converting the mismatches in the stem and the unpaired bases in the loop to well-matched base pairs in duplex provides the enthalpic driving force. The values for ΔG also show an oscillating pattern depending on whether n = even or n = odd, although the difference is

much less pronounced than for ΔH and $T\Delta S$. Importantly, as n increases, the duplex is increasingly more thermodynamically stable than the stem-loop hairpins.

Table 3.4 Thermodynamic Parameters for Conversion of $(CAG)_n$ and $(CTG)_n$ Stem-Loop Hairpins to $(CAG)_n/(CTG)_n$ Duplex.^{1,2}

| n | ΔH (kcal mol ⁻¹) | $T\Delta S$ ³ (kcal mol ⁻¹) | ΔG ³ (kcal mol ⁻¹) |
|----|---|---|--|
| 6 | -59.4 ± 3.2 | -51.2 ± 2.8 | -8.2 ± 0.4 |
| 7 | -86.0 ± 2.8 | -73.6 ± 2.5 | -12.4 ± 0.3 |
| 8 | -80.7 ± 2.3 | -66.7 ± 2.1 | -13.9 ± 0.2 |
| 9 | -113 ± 4 | -94.5 ± 3.2 | -18.7 ± 0.5 |
| 10 | -116 ± 5 | -94.0 ± 3.9 | -21.8 ± 1.5 |
| 11 | -135 ± 3 | -109 ± 3 | -25.9 ± 2.2 |
| 12 | -138 ± 7 | -109 ± 6 | -28.1 ± 1.0 |
| 13 | -163 ± 5 | -132 ± 5 | -31.3 ± 0.7 |
| 14 | -154 ± 8 | -121 ± 7 | -32.7 ± 1.2 |

¹ In 20 mM sodium phosphate, 100 mM NaCl, pH 7.0.

² Errors represent standard deviation from the analysis of three scans of a single sample preparation.

³ Values at 37 °C.

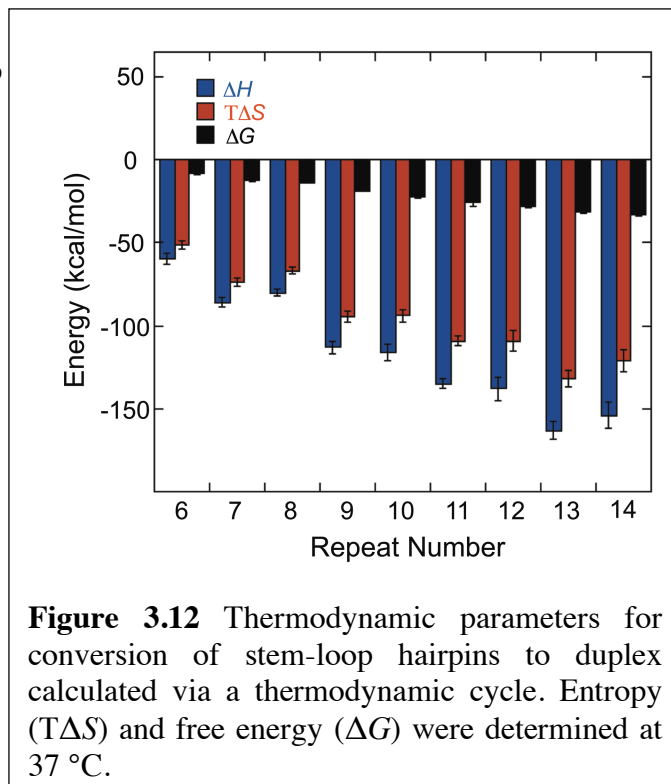


Figure 3.12 Thermodynamic parameters for conversion of stem-loop hairpins to duplex calculated via a thermodynamic cycle. Entropy ($T\Delta S$) and free energy (ΔG) were determined at 37 °C.

In light of our observation that the canonical duplex is increasingly thermodynamically favored over stem-loop hairpins as the repeat length increases, the question arises whether this conversion will proceed within a biologically-relevant time scale. Therefore, we also characterized the kinetic parameters describing the stem-loop hairpin to duplex conversion.

3.4.6 Kinetic Analysis of Stem-Loop Hairpin to Duplex Conversion

In order to determine how repeat number influences the rate at which non-canonical stem-loop hairpins convert to duplex, we performed a spectroscopy-based kinetic analysis. In these experiments, complementary $(CAG)_n$ and $(CTG)_n$ stem-loop hairpins were combined at 37 °C and

the absorbance at 260 nm was monitored as a function of time. The data were fit to a second-order equation to obtain the rate of stem-loop hairpin to duplex conversion (Table 3.5).²² Importantly, at the DNA concentrations used for these kinetic analyses, no homoduplex was observed by native PAGE, and the starting species were exclusively stem-loop hairpins (data not shown). Over the range of $n = 6-14$ we observe that when n is odd, the stem-loop hairpins convert to duplex faster than when n is even (Figure 3.13). Furthermore, when $n = \text{even}$ there is a statistically significant decrease from $n = 6$ to $n = 8$, but with continued increase in n , the rate remains the same (Figure 3.13A), when $n = \text{odd}$ the rate of conversion to duplex decreases with n (Figure 3.13B). To determine whether the rates of hairpin to duplex conversion reach a lower limit with increasing repeats, especially when $n = \text{even}$, we measured the rate at which $(\text{CAG})_{25}$ and $(\text{CTG})_{25}$, and also $(\text{CAG})_{30}$ and $(\text{CTG})_{30}$, stem-loop hairpins convert to duplexes. When $n = 25$, a significant decrease in rate for hairpin to duplex conversion is observed relative to $n = 7, 9, 11, 13$; however when $n = 30$, only a small decrease in rate was observed relative to when $n = 8, 10, 12$, or 14 (Table 3.5, Figure 3.13).

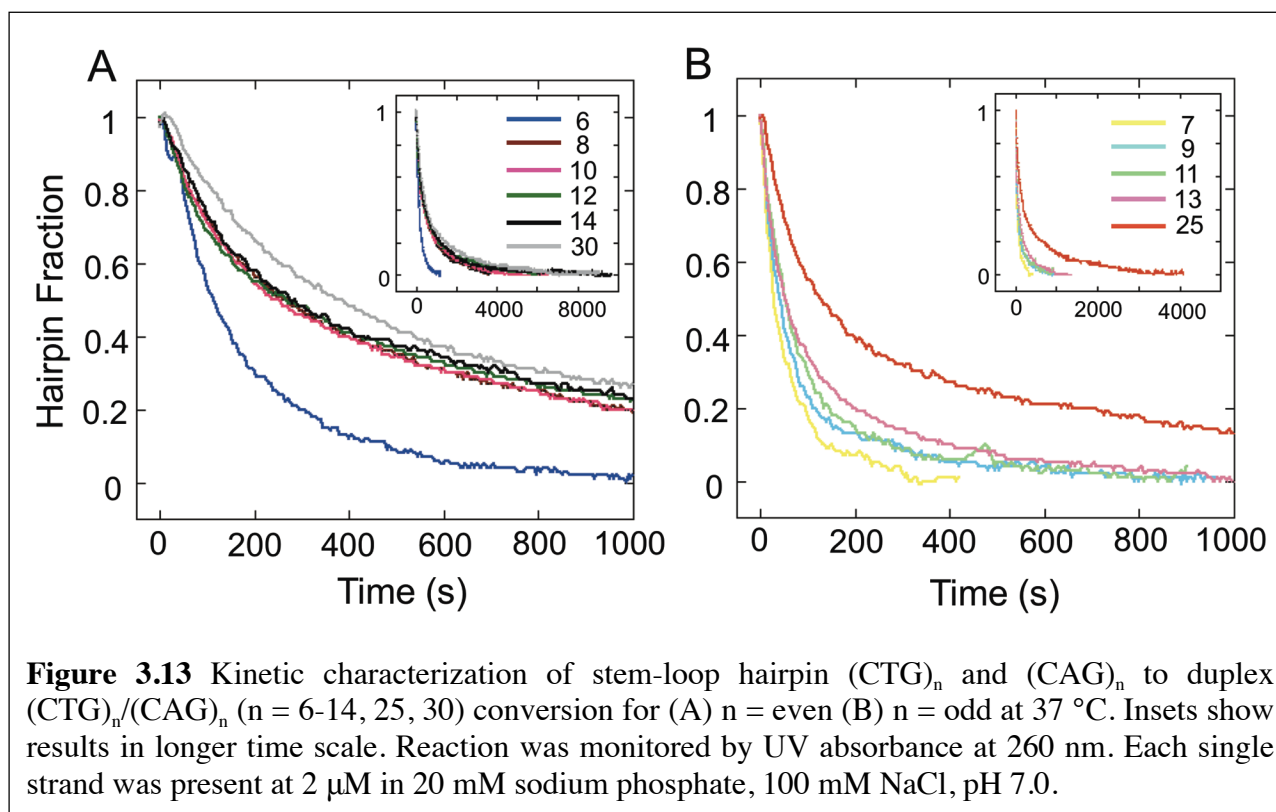
Table 3.5 Rate and Activation Enthalpy for Stem-Loop Hairpin to Duplex Conversion at 37 °C.¹

| n | $k_{37^\circ\text{C}}^2 (\text{M}^{-1} \text{s}^{-1}) / \times 10^3$ | $\Delta H^\ddagger^3 (\text{kcal mol}^{-1})$ |
|-----|--|--|
| 6 | 15.2 ± 2.0 | 29.1 ± 3.2 |
| 7 | 44.4 ± 2.7 | 6.76 ± 0.57 |
| 8 | 6.2 ± 0.6 | - |
| 9 | 40.2 ± 1.6 | - |
| 10 | 5.7 ± 0.4 | 35.8 ± 4.1 |
| 11 | 31.4 ± 2.3 | 21.4 ± 3.6 |
| 12 | 5.1 ± 0.9 | - |
| 13 | 26.8 ± 1.7 | - |
| 14 | 6.3 ± 0.8 | 41.0 ± 5.1 |
| 25 | 11.0 ± 0.5 | - |
| 30 | 4.6 ± 0.2 | - |

¹ In 20 mM sodium phosphate, 100 mM NaCl, pH 7.0.

² Errors represent standard deviation from a minimum of three experiments.

³ Errors represent the uncertainty derived from fitting the rate constants at four different temperatures to the Eyring equation.



The conversion from stem-loop hairpins to duplex is proposed to involve two possible mechanisms: interaction of the complementary stem-loop hairpins by either the loop regions or the stem termini.^{27,32} We previously showed that both mechanisms are utilized when $(CAG)_{10}$ and $(CTG)_{10}$ stem-loop hairpins convert to duplex.³² We also previously reported that the rate of formation of $(CAG)_{10}/(CTG)_{11}$ duplex is 4-fold faster than formation of the $(CAG)_{10}/(CTG)_{10}$ duplex.¹⁶ We postulated that the unpaired nucleobases of the overhanging repeat of the $(CTG)_{11}$ stem-loop hairpin facilitated interactions at the termini of the stems and this additional triplet overhang is responsible for the faster conversion to duplex. In the current work, we show that when both stem-loop hairpins have $n = \text{odd}$, the rate of conversion to duplex is faster than when $n = \text{even}$. We propose that this faster rate of duplex formation is aided by the presence of the overhangs on the stem-loop hairpins.

By conducting similar kinetics experiments as a function of temperature, we determined the activation enthalpy (ΔH^\ddagger), which is used to characterize the transition state during the process of converting from the non-canonical stem-loop hairpins to canonical duplex (Figure 3.14). The error associated with this experiment is large; thus, ΔH^\ddagger are not reported for all values of n , but rather for $n = 6, 7, 10, 11$ and 14 (Table 3.5). For both $n = \text{even}$ and $n = \text{odd}$, the values of ΔH^\ddagger increase as a function of n . This result

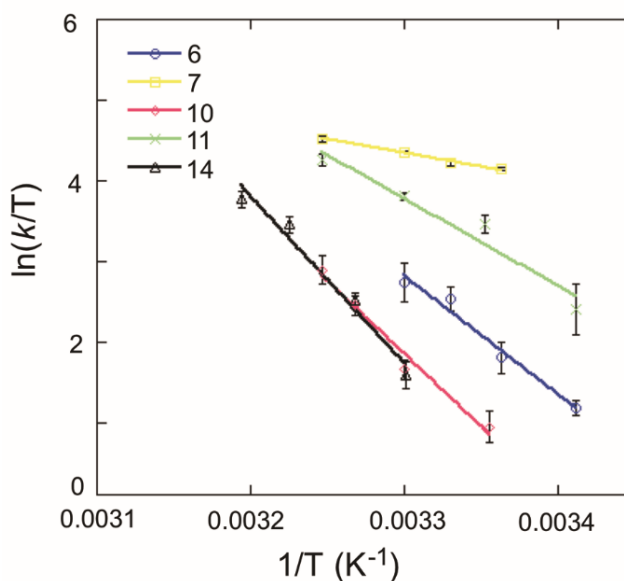


Figure 3.14 Determining the ΔH^\ddagger by Eyring equation for $(CTG)_n$ and $(CAG)_n$ stem-loop hairpin to $(CAG)_n/(CTG)_n$ duplex conversion for $n = 6, 7, 10, 11, 14$. Each data point represents an average of at least three experiments performed at a given temperature. Lines represent the least square linear fits. Temperatures were chosen below the melting temperature of the stem-loop hairpins.

is consistent with the rates of stem-loop hairpin to duplex conversion obtained at 37 °C where the rates *decrease* as a function of n both when $n = \text{even}$ and $n = \text{odd}$. Indeed, this increase in ΔH^\ddagger provides a potential rationale for the propensity of longer TNR sequences to expand – the non-canonical stem-loop hairpins persist rather than convert to canonical duplex. Besides, the ΔH^\ddagger determined for each hairpin-to-duplex conversion is less than the total enthalpy required for melting the $(CTG)_n$ and $(CAG)_n$ stem-loop hairpins. This observation confirms our previous proposal that the conversion from stem-loop hairpins to canonical duplex does not require the stem-loop hairpins to be denatured globally;^{31,33} rather, the structured stem-loop hairpins interact via the loop and/or stem regions to convert to duplex.

3.5 Conclusion

In summary, our results reveal differences for both $(CAG)_n$ and $(CTG)_n$ stem-loop hairpins depending on whether $n = \text{even}$ or $n = \text{odd}$; differences in the thermodynamic stability of the non-canonical stem-loop hairpins, rate of conversion to canonical duplex, and ΔH^\ddagger were observed (Figure 3.15A). The conversion of hairpin to duplex is thermodynamically and kinetically more

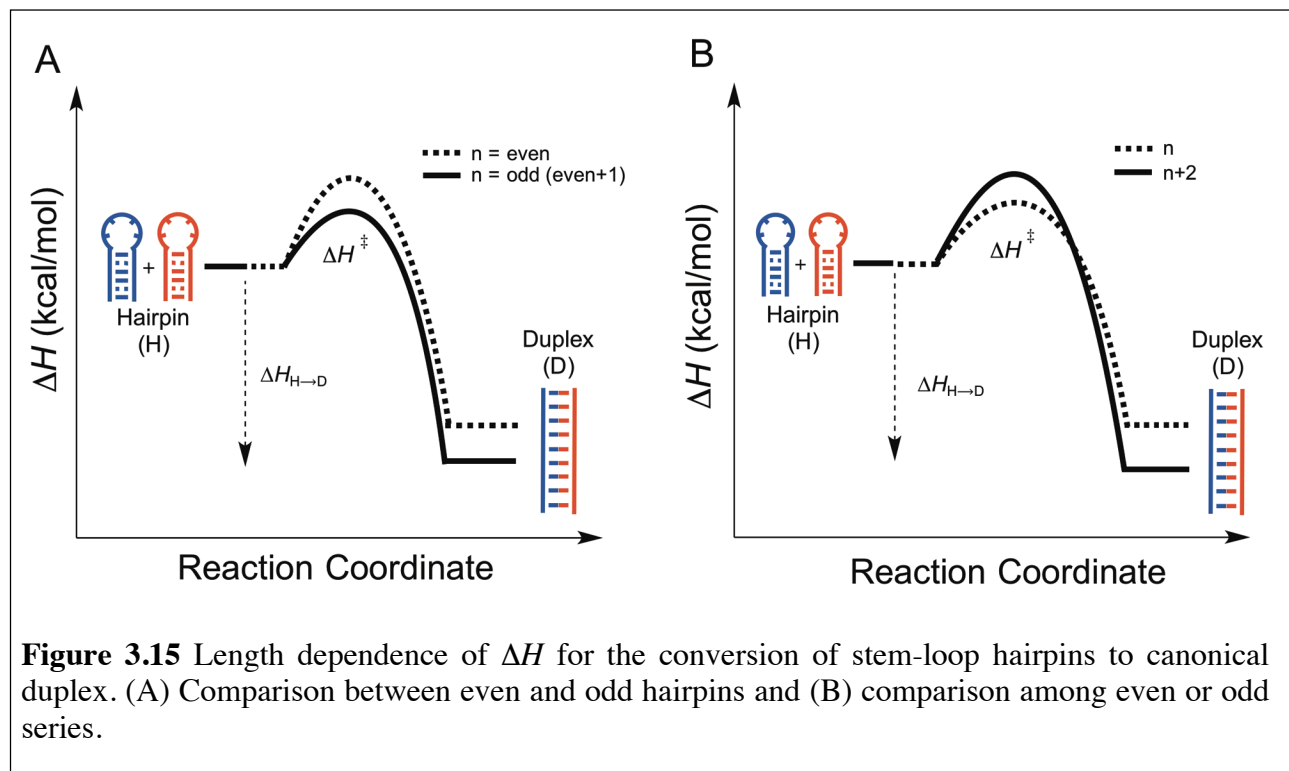


Figure 3.15 Length dependence of ΔH for the conversion of stem-loop hairpins to canonical duplex. (A) Comparison between even and odd hairpins and (B) comparison among even or odd series.

favorable for $n = \text{odd}$, as indicated by the more stabilized duplex $\Delta H_{H \rightarrow D}$ and lower activation enthalpy ΔH^\ddagger when compared to $n = \text{even}$ (Figure 3.15A). As a result of the even/odd pattern, changing the length of hairpins does not linearly influence the biophysical properties of either stem-loop hairpin unfolding or the hairpin-to-duplex conversion. However, within the even ($n = 6, 8, 10, 12, 14$) or odd ($n = 7, 9, 11, 13$) series, in other words, by increasing the repeat number n by 2, the hairpin-to-duplex conversion is thermodynamically more favorable but kinetically unfavorable as evidenced by the more stabilized duplex $\Delta H_{H \rightarrow D}$ but higher ΔH^\ddagger (Figure 3.15B).

Taken together, our analyses of length dependence of (CTG)_n and (CAG)_n hairpin related thermodynamic and kinetic properties may contribute to the understanding of the role of non-canonical structures in TNR expansion and should be considered when evaluating proposed mechanisms for expansion.

3.6 References

- (1) McMurray, C. T. Mechanisms of Trinucleotide Repeat Instability during Human Development. *Nat. Rev. Genet.* **2010**, *11*, 786–799.
- (2) Kozlowski, P.; de Mezer, M.; Krzyzosiak, W. J. Trinucleotide Repeats in Human Genome and Exome. *Nucleic Acids Res.* **2010**, *38*, 4027–4039.
- (3) Lopez Castel, A.; Cleary, J. D.; Pearson, C. E. Repeat Instability as the Basis for Human Diseases and as a Potential Target for Therapy. *Nat. Rev. Mol. Cell. Biol.* **2010**, *11*, 165–170.
- (4) Mirkin, S. M. Expandable DNA Repeats and Human Disease. *Nature* **2007**, *447*, 932–940.
- (5) Huntington's Disease Collaborative Research Group. A Novel Gene Containing a Trinucleotide Repeat That Is Expanded and Unstable on Huntington's Disease Chromosomes. *Cell* **1993**, *72*, 971–983.
- (6) Kang, S.; Jaworski, A.; Ohshima, K.; Wells, R. D. Expansion and Deletion of CTG Repeats from Human Disease Genes Are Determined by the Direction of Replication in *E. coli*. *Nat. Genet.* **1995**, *10*, 213–218.
- (7) Kang, S.; Ohshima, K.; Shimizu, M.; Amirhaeri, S.; Wells, R. D. Pausing of DNA Synthesis *in vitro* at Specific Loci in CTG and CGG Triplet Repeats from Human Hereditary Disease Genes. *J. Biol. Chem.* **1995**, *270*, 27014–27021.

- (8) Samadashwily, G. M.; Raca, G.; Mirkin, S. M. Trinucleotide Repeats Affect DNA Replication *in vivo*. *Nat. Genet.* **1997**, *17*, 298–304.
- (9) Fouche, N.; Ozgur, S.; Roy, D.; Griffith, J. D. Replication Fork Regression in Repetitive DNAs. *Nucleic Acids Res.* **2006**, *34*, 6044–6050.
- (10) Kovtun, I. V.; Liu, Y.; Bjoras, M.; Klungland, A.; Wilson, S. H.; McMurray, C. T. OGG1 Initiates Age-Dependent CAG Trinucleotide Expansion in Somatic Cells. *Nature* **2007**, *447*, 447–452.
- (11) Jarem, D. A.; Wilson, N. R.; Schermerhorn, K. M.; Delaney, S. Incidence and Persistence of 8-Oxo-7,8-Dihydroguanine within a Hairpin Intermediate Exacerbates a Toxic Oxidation Cycle Associated with Trinucleotide Repeat Expansion. *DNA Repair* **2011**, *10*, 887–896.
- (12) McMurray, C. T. Hijacking of the Mismatch Repair System to Cause CAG Expansion and Cell Death in Neurodegenerative Disease. *DNA Repair* **2008**, *7*, 1121–1134.
- (13) Paiva, A. M.; Sheardy, R. D. Influence of Sequence Context and Length on the Structure and Stability of Triplet Repeat DNA Oligomers. *Biochemistry* **2004**, *43*, 14218–14227.
- (14) Amrane, S.; Mergny, J.-L. Length and pH-Dependent Energetics of (CCG)_n and (CGG)_n Trinucleotide Repeats. *Biochimie* **2006**, *88*, 1125–1134.
- (15) Amrane, S. Length-Dependent Energetics of (CTG)_n and (CAG)_n Trinucleotide Repeats. *Nucleic Acids Res.* **2005**, *33*, 4065–4077.
- (16) Figueroa, A. A.; Cattie, D.; Delaney, S. Structure of Even/odd Trinucleotide Repeat Sequences Modulates Persistence of Non-B Conformations and Conversion to Duplex. *Biochemistry* **2011**, *50*, 4441–4450.
- (17) Marky, L. a; Breslauer, K. J. Calculating Thermodynamic Data for Transitions of Any Molecularity from Equilibrium Melting Curves. *Biopolymers* **1987**, *26*, 1601–1620.

- (18) Chalikian, T. V.; Völker, J.; Plum, G. E.; Breslauer, K. J. A More Unified Picture for the Thermodynamics of Nucleic Acid Duplex Melting: A Characterization by Calorimetric and Volumetric Techniques. *Proc. Natl. Acad. Sci. U. S. A.* **1999**, *96*, 7853–7858.
- (19) Jarem, D. a.; Wilson, N. R.; Delaney, S. Structure-Dependent DNA Damage and Repair in a Trinucleotide Repeat Sequence. *Biochemistry* **2009**, *48*, 6655–6663.
- (20) Warshaw, M. M.; Tinoco, I. Optical Properties of Sixteen Dinucleoside Phosphates. *J. Mol. Biol.* **1966**, *20*, 29–38.
- (21) Gacy, a M.; McMurray, C. T. Influence of Hairpins on Template Reannealing at Trinucleotide Repeat Duplexes: A Model for Slipped DNA. *Biochemistry* **1998**, *37*, 9426–9434.
- (22) Gacy, a M.; McMurray, C. T. Hairpin Formation within the Human Enkephalin Enhancer Region. 1. Kinetic Analysis. *Biochemistry* **1994**, *33*, 11951–11959.
- (23) Ramalanjaona, N.; Rocquigny, H. De; Millet, A.; Ficheux, D.; Darlix, J.-L.; Mély, Y. Investigating the Mechanism of the Nucleocapsid Protein Chaperoning of the Second Strand Transfer during HIV-1 DNA Synthesis. *J. Mol. Biol.* **2007**, *374*, 1041–1053.
- (24) Vander Meulen, K. A.; Butcher, S. E. Characterization of the Kinetic and Thermodynamic Landscape of RNA Folding Using a Novel Application of Isothermal Titration Calorimetry. *Nucleic Acids Res.* **2012**, *40*, 2140–2151.
- (25) Chen, C.; Wang, W.; Ge, J.; Zhao, X. S. Kinetics and Thermodynamics of DNA Hybridization on Gold Nanoparticles. *Nucleic Acids Res.* **2009**, *37*, 3756–3765.
- (26) Chi, L. M.; Lam, S. L. Structural Roles of CTG Repeats in Slippage Expansion during DNA Replication. *Nucleic Acids Res.* **2005**, *33*, 1604–1617.
- (27) Bernacchi, S.; Ennifar, E.; Tóth, K.; Walter, P.; Langowski, J.; Dumas, P. Mechanism of

- Hairpin-Duplex Conversion for the HIV-1 Dimerization Initiation Site. *J. Biol. Chem.* **2005**, *280*, 40112–40121.
- (28) Marky, L. A.; Breslauer, K. J. Calculating Thermodynamic Data for Transitions of Any Molecularity from Equilibrium Melting Curves. *Biopolymers* **1987**, *26*, 1601-1620
- (29) Mitas, M. Trinucleotide Repeats Associated with Human Disease. *Nucleic Acids Res.* **1997**, *25*, 2245–2254.
- (30) Arnold, F. H.; Wolk, S.; Cruz, P.; Tinoco, I. Structure, Dynamics, and Thermodynamics of Mismatched DNA Oligonucleotide Duplexes d(CCCAGGG)₂ and d(CCCTGGG)₂. *Biochemistry* **1987**, *26*, 4068–4075.
- (31) Volker, J.; Makube, N.; Plum, G. E.; Klump, H. H.; Breslauer, K. J. Conformational Energetics of Stable and Metastable States Formed by DNA Triplet Repeat Oligonucleotides: Implications for Triplet Expansion Diseases. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 14700–14705.
- (32) Avila Figueroa, A.; Delaney, S. Mechanistic Studies of Hairpin to Duplex Conversion for Trinucleotide Repeat Sequences. *J. Biol. Chem.* **2010**, *285*, 14648–14657.
- (33) Paiva, A. M.; Sheardy, R. D. The Influence of Sequence Context and Length on the Kinetics of DNA Duplex Formation from Complementary Hairpins Possessing (CNG) Repeats. *J. Am. Chem. Soc.* **2005**, *127*, 5581–5585.

Chapter 4: Differential Scanning Calorimetry Study of

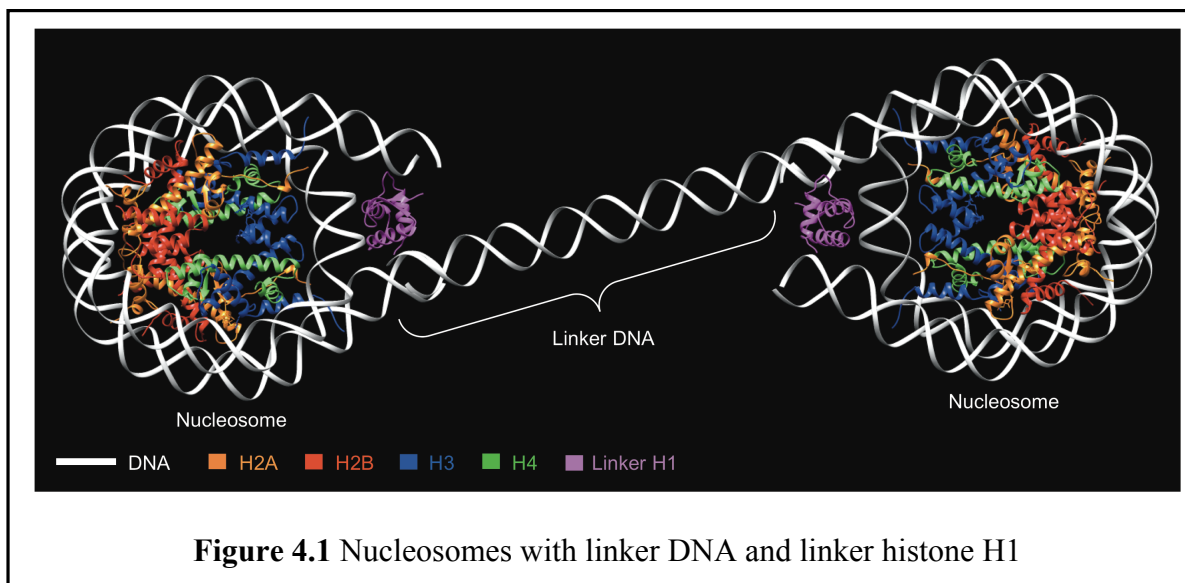
A Nucleosome and its Subunits

4.1 Abstract

DNA in eukaryotic cells is packaged around histone proteins and forms the basic repeating unit of chromatin as nucleosome. In this work, we use differential scanning calorimetry (DSC) to characterize the thermodynamics of the nucleosome and its subunits (DNA and histone core). The results from free nucleosomal DNA analysis show that the melting of the DNA is highly sequence-dependent and that multiple transitions may occur from unfolding different domains of the DNA. The melting of nucleosomal DNA is also ion strength-dependent: 0.1 M salt is required for proper stabilization of the DNA. Results from analyzing the histone protein components show that H2A/H2B dimer is more stable than (H3-H4)₂ tetramer, with only H2A/H2B exhibiting reversible thermal unfolding under low ion strength conditions. Unfolding of histone octamer is consistent with the results from its dimer and tetramer, with (H3-H4)₂ tetramer unfolding at lower temperature than the H2A/H2B dimer. In addition, unfolding of tetramer is accompanied by its dissociation from the two H2A/H2B dimers. DSC analysis of nucleosomes suggests that the nucleosomal DNA unwraps from the histone core at around 72 °C and that the DNA sequence effect is minimal. This result is consistent with the thermal analysis of higher order DNA structure in nuclei or whole cells, and provides insight regarding the long-running controversy in assigning the DNA conformational change during the thermal denaturation process. The study will also provide information about nucleosome stability and the contributing factors, which influence DNA replication, transcription, repair and recombination.

4.2 Introduction

The nucleosome is the fundamental repeating unit of chromatin in eukaryotic cells. It is composed of around 146 bp of nucleosomal DNA wrapped in 1.7 left-handed superhelical turns around a histone core consisting of two copies of four histone proteins, H2A, H2B, H3 and H4.^{1,2} Nucleosomes are connected by linker DNA (typically 10-80 bp)³ and linker histone H1 to form the high order structures (Figure 4.1).⁴ Wrapped nucleosomal DNA is potentially much less accessible to cellular processes involving DNA as the substrate for several possible reasons. For example, nucleosomal DNA is sterically occluded due to its close interaction with the histone protein core and also the other superhelical turns of DNA within the same nucleosome. It is therefore likely that during the process of DNA replication, transcription, recombination and DNA repair, DNA is transiently relaxed from the histone core to allow various proteins to access the DNA within the nucleosome.⁵ Since most DNA exists as the form of nucleosomes, it is of great interest to understand the stability of the nucleosome and how different components affect the unfolding process of nucleosomes.



Traditional measurement of DNA thermal stability by UV-vis follows the hyperchromic change that occur as DNA melts, and requires clear solution.⁶ Chromatin is soluble only in low salt conditions but will aggregate at physiological salt range, which makes it impossible to analyze by spectrometric-based methods due to scatter from the aggregated particles. For this reason, differential scanning calorimetry (DSC) has been widely used for the thermal characterization of chromatin DNA and more complicated biological systems under physiologically relevant conditions. In the early 1980s, isolated chromatin, intact nuclei, or even the whole cells, have been characterized by DSC to probe the DNA structural changes induced by temperature. In these studies, the roles of ionic strength, pH, and also the presence of the H1 histone were investigated.⁶⁻¹³ DSC has also been employed to monitor structural changes in chromatin during the cell aging process or when treated by various chemotherapeutic agents.^{14, 15} Although these DSC studies provided a broad perspective of DNA thermal stabilities, it is quite challenging to assign individual transitions to the corresponding chromatin DNA unfolding process. In fact, contradictory results have been reported from different research groups even though the thermal denaturation profiles are consistent.^{8, 9} Furthermore, the coupling of post-translational modification of histone proteins and DNA sequence heterogeneity makes it difficult to provide detailed information about the factors that contribute to the stability of DNA in chromatin when samples are prepared from biological sources.^{5, 16} Therefore, the study of a simple control system, such as the nucleosome and its subunits, has become a focus in recent years. Several studies have examined the stabilities of the histone H2A/H2B dimer and the (H3-H4)₂ tetramer by spectroscopic- or DSC-based analysis.¹⁷⁻¹⁹ Besides the histone protein components, DNA sequences have also been investigated for their effects on the nucleosome stability. Several natural or artificial nucleosomal DNA

sequences have been discovered that show strong affinity towards histone protein core and only adopt a single translational position relative to the histone octamer.²⁰⁻²²

In this study, we constructed a model nucleosome system using a bottom-up approach with recombinant histone proteins. The thermal stabilities of histone H2A/H2B dimer, (H3-H4)₂ tetramer and octamer were examined under various salt conditions by DSC. We also employed two model nucleosome positioning DNA sequences (S1 and 601 DNA), which only adopt a single translational position relative to the histone core, to evaluate the sequence effect on nucleosome stability by DSC. In particular, as discussed in Chapter 3, we want to explore the thermodynamic consequences of by trinucleotide (TNR) DNA expansion in Huntington's disease (HD). Clinical data from HD patients showed a narrow range of CAG/CTG repeat number variation from healthy (5-35) to diseased state (>40).²³ Besides, an earlier onset of disease and severe symptoms are anticipated when patients have a larger tract of CAG/CTG repeats.²³ It is therefore interesting to know how the variation of CAG/CTG repeat number contributes to the propensity of TNR expansion in the nucleosome level. Hence, the S1 sequence with 10, 20 and 30 of CAG repeats were prepared, representing the progression of DNA expansion in HD, and the effect of repeat number on the nucleosome stability is evaluated by DSC. Results from these analyses provide insight into the understanding of DNA stability in chromatin and potential mechanistic details of the repeat length-dependent TNR expansion.

4.3 Experimental Procedure

4.3.1 Preparation of Nucleosomal DNA

All nucleosomal DNA template (601, S1, S1-10CAG, S1-20CAG, S1-30CAG) sequences were synthesized and purified according to methods published previously in our laboratory²⁴ and

Table 4.1 DNA sequences used in this study

| DNA | Sequence (all sequences are shown from 5' to 3' direction) |
|-------------|---|
| S1* | ATCAATTAATAGTTGAAGTTGTAGTAAATGTTAATGTAGGTCTGTTGTTCCGAT ATTACCAAACCTTCACCTAATAGCGTAGTACACAGGAGAAGGACATGAACA TGAACCTAATGAACACAACAAATAATGTAAGTGCCCGAT |
| S1-10CAG* | ATCAATTAATAGTTGAAGTTGTAGTAAATGTTAATGTAGGTCTGTTGTTCCGAT ATTACCACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGGAGAAGGACATGAAC ATGAACCTAATGAACACAACAAATAATGTAAGTGCCCGAT |
| S1-20CAG* | ATCAATTAATAGTTGAAGTTGTAGTAAATGTTAATGTAGGTCTGTTCCAGCAGCA GCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGGAGCAGCAGCAGCAG CAGAACCTAATGAACACAACAAATAATGTAAGTGCCCGAT |
| S1-30CAG* | ATCAATTAATAGTTGAAGTTGTAGTAAATGTCAGCAGCAGCAGCAGCAGCAGCAGC AGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCA GCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCA |
| 601* | ATCGATGTATATATCTGACACGTGCTTGGAGACTAGGGAGTAATCCCCTTGCGC GTTAAAACGCGGGGGACAGCGGTACGTGCGTTTAAGCGGTGCTAGAGCTGTC TACGACCAATTGAGCGGCCTCGGCACCGGGATTCTGAT |
| S1 primers | Forward: TGACTAGAATTCTGACTAGCTAGCTGATATCAATTAATAGTTGAAGTT GTAG Reverse: TAGTCATCTAGATAATTAGATATCGGGCACTTACATTAT |
| 601 primers | Forward: TGACTAGAATTCTGACTAGCTAGCTGATATCGATGTATATATCTGACA CGTGCTTGGAGACTAG Reverse: TAGTCATCTAGATAATTAGATATCAGAATCCCGGTGCCGAG |

* Sequence is only shown for one strand of the nucleosomal DNA

the actual template sequences are listed in Table 4.1. In order to generate large scale of nucleosomal DNA, the above DNA templates were cloned into pUC19 plasmid according to published strategies to generate plasmids contain multiple repeats of give nucleosomal DNA sequences.²⁵⁻²⁷ In general, template DNA was PCR amplified with forward and reverse primers to introduce one EcoRI, one XbaI, one NheI and two EcoRV restriction sites. The PCR product DNA were double digested with EcoRI and XbaI, the fragment containing DNA of interest was purified and ligated into the pUC19 EcoRI and XbaI sites, followed by transformation and amplification in DH5 α cells. Small amount of the obtained plasmids were double digested by EcoRI and NheI to generate the plasmid backbone for next round ligation, the rest of the plasmids were double digested by EcoRI and XbaI to generate the insert and was ligated into the EcoRI and NheI sites prepared above. The process is continued until the plasmid contains 16 repeats of sequence of interest. The prepared plasmids were then amplified in large scale in 2 L of cell culture and

plasmids were purified according to protocol.²⁷ The obtained plasmids were then digested with EcoRV to retrieve the nucleosomal DNA. After ethanol precipitation, the nucleosomal DNA were further purified by HPLC on a DEAE column () to remove the small fragment between the two tandem repeats. The DNA collected from HPLC were concentrated and store at -20 °C.

4.3.2 Histone Expression and Purification

Individual full length histone proteins from *Xenopus laevis* were expressed and purified according to published protocol²⁸ with several modifications. In short, BL21(DE3) cells that expressed histone proteins were lysed in 50 mM Tris-HCl, 100 mM NaCl, 1 mM EDTA, 1 mM benzamidine, 5 mM 2-mercaptoethanol, pH 7.5 buffer by several rounds of “freeze-and-thaw cycle”, followed by homogenization on ice to reduce the viscosity. Inclusion bodies were collected by centrifugation and purified with four washing steps with lysis buffer that supplemented with 1% Triton X-100 (v/v) or 2 M NaCl in the first three washes. Histones were then extracted from inclusion bodies with DMSO and unfolding buffer (7 M guanidinium-HCl, 20 mM Tris-HCl, 10 mM DTT, pH 7.5). The extracted histones were loaded into Sephacryl S-200 high resolution size exclusion column (2.5 cm ID × 75 cm height) to remove the sheared *E. coli* genomic DNA under unfolding conditions (7 M urea (batch deionized), 20 mM sodium acetate, 1 M NaCl, 1 mM EDTA, 5 mM 2-mercaptoethanol, pH 5.2). The collected fractions were then dialyzed against H₂O with 2 mM 2-mercaptoethanol and dried by lyophilization. The obtained histones were then purified by both anion- and cation-exchang chromatography with salt gradient (200 mM to 1 M NaCl) under denatured conditions. After the gradient, columns were regenerated with unfolding buffer (the majority of H3 histone will elute in this regeneration step while all other three histones elute within

the salt gradient). The purified histones were dialysis against H₂O with 2 mM 2-mercaptoethanol and dried in aliquots by lyophilization.

4.3.3 Histone H2A/H2B Dimer, (H3-H4)₂ Tetramer and Histone Octamer Assembly

Histone H2A/H2B dimer, (H3-H4)₂ tetramer and histone octamer were assembled according to Luger and coworkers.²⁸ Lyophilized histone proteins (~6-8 mg each) were resuspended in unfolding buffer and the concentrations were determined by UV-vis at 276 nm. Equal molars of each histone were mixed together and diluted with unfolding buffer to 1 mg/mL. The mixture was then dialyzed against refolding buffer (2 M NaCl, 10 mM Tris-HCl, 1 mM EDTA, 5 mM 2-mercaptoethanol, pH 7.5) overnight at 4 °C with three changes of buffer. The obtained solution was concentrated to ~1 mL and loaded into a Superdex 200 size exclusion column (1.5 cm ID X 70 cm height). Fractions were collected and monitored by UV-vis at 276 nm, all fractions absorbed at 276 nm were characterized by SDS-PAGE. Fractions contain all four histone proteins were combined and concentrated, adjusted with glycerol and store at -20 °C.

4.3.4 Nucleosome Reconstitution and Purification

Histone octamer was first dialyzed against refolding buffer overnight to remove glycerol and quantitated with UV-vis afterwards. Nucleosomal DNA was adjusted with 4 M KCl to make the final salt concentration 2 M and add DTT to a final concentration of 10 mM. DNA was then incubated at 4 °C for 30 min. Histone octamer was then added to DNA with a molar ratio 1 : 1.1 with slightly excess of DNA. The mixture was then dialyzed against 2 M KCl with continuous reducing salt concentration overnight described in the published protocol.²⁸ After the salt gradient is finished, the reconstituted nucleosome was dialyzed in 10 mM Tris-HCl, 100 mM KCl, 1 mM

DTT, 1 mM EDTA, pH 7.5 for 3 hours. For nucleosomes reconstituted with S1 series DNA, the obtained solutions were incubated at 55 °C for 1 hour, while this step is not necessary for nucleosome reconstituted with 601 DNA. Nucleosomes were then purified by HPLC with the DEAE column at 10 °C and the same salt gradient mentioned above.

4.3.5 DSC Characterization of Nucleosomes

After the HPLC purification, nucleosomes were immediately buffer exchanged to DSC buffer (20 mM KPi, 100 mM KCl pH 7.2) using the 3 K MWCO concentrators. The concentrations were determined by UV-vis at 260 nm. nucleosome solutions were prepared for DSC experiments by diluting the stock purified nucleosomes in DSC buffer to desired concentration. 10 uL of the degassed nucleosome solution was analyzed by DEAE column on HPLC to characterize the sample loaded into the DSC.

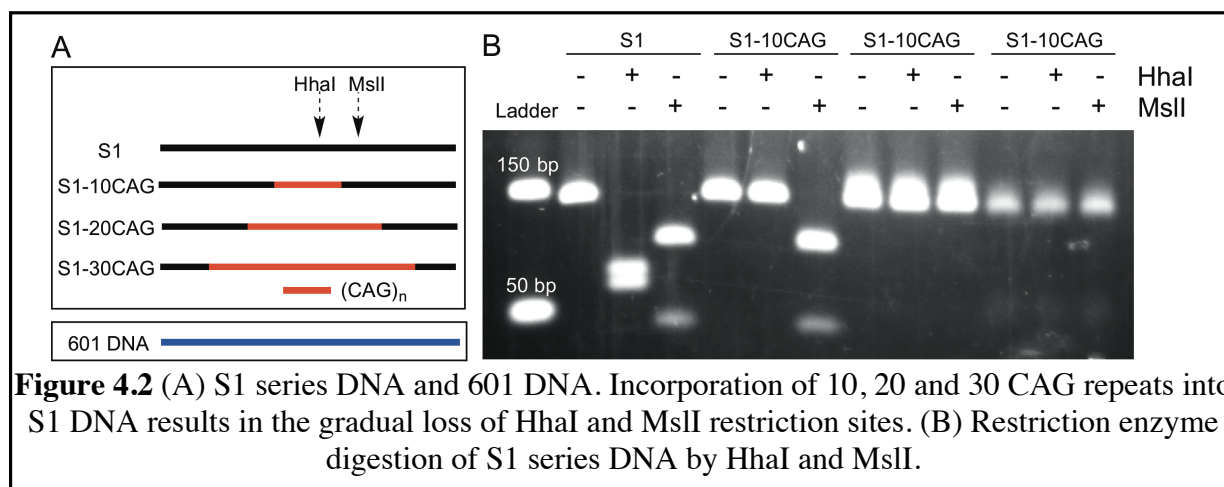
Calorimetric experiments were performed using a TA Instruments Nano DSC III. The S1 series nucleosomes were at concentrations of 100 μ M and the 601 nucleosome was at concentrations of 50 μ M. All nucleosomes were suspended in 20 mM potassium phosphate, 100 mM KCl, pH 7.2. Both the oligonucleotide samples and reference buffer sample were degassed *in vacuo* for 15 min at 25 °C before analysis. All data were recorded with TA Instrument Nano DSCRun software version 4.2.6. Data were obtained by continuously monitoring the excess power required to maintain both sample cell and reference cell at the same temperature. The samples were heated from 4 °C to 100 °C followed by cooling from 100 °C to 4 °C, both at 1.0 °C/min. The sample equilibrated for 10 min at 4 and 100 °C between each cooling and heating cycle, respectively. The resulting thermograms display excess heat capacity as a function of temperature. A buffer reference was analyzed using the same procedure described above and the thermograms

were corrected using this background. Further analyses were performed on TA NanoAnalyze software version 3.5.1. The melting temperature (T_m) was determined as the T_{max} of each transition.

4.4 Results and Discussion

4.4.1 Preparation of Nucleosomal DNA

In order to evaluate the sequence effect on the thermal stability of nucleosome, two nucleosome positioning sequences, S1 and 601 DNA, were prepared in this study. 601 DNA is by far the strongest nucleosome positioning sequence screened by SELEX while S1 DNA is a moderately strong positioning sequence, which has been used previously in our lab for nucleosome reconstitutions.^{21, 24, 29, 30} Both sequences adopt a single translational position relative to the histone octamer core to obtain homogeneous nucleosome. Additionally, in order to systematically evaluate how the number of CAG repeats affects nucleosome incorporation and nucleosome stability, the central region of the S1 sequence was replaced by either 10, 20, or 30 CAG repeats (defined as S1-10CAG, S1-20CAG, S1-30CAG) while maintaining the flanking sequences on both sides. In this way, any observed changes will be contributed from the variation in the CAG repeat number (Figure 4.2A).

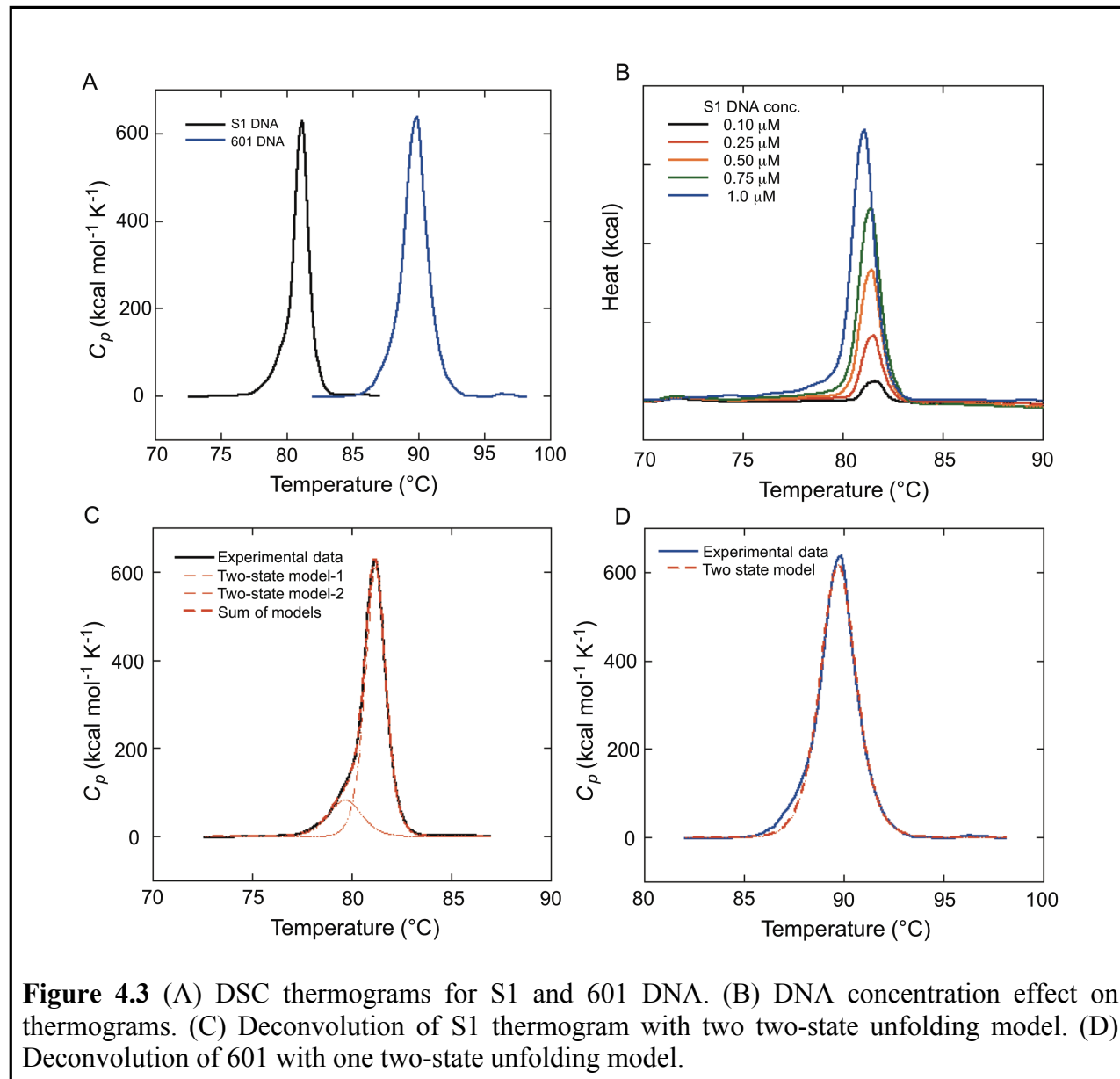


The generation of each sequence in the S1 series DNA involves the same primers, and the PCR reactions become increasingly difficult when the template DNA includes more CAG repeats. Therefore, it is very important that there is no cross-contamination between DNA templates. Our results show that for any CAG repeat-containing template DNA, even trace amount of S1 template contamination will dominate the PCR product. This result is consistent with previous reports that polymerase stalls at TNR sites during replication.^{31,32} For 601 and S1 DNA, the sequence identities are confirmed by Maxam-Gilbert DNA sequencing. The successful incorporation of 10, 20 and 30 CAG repeats in S1 DNA is confirmed by both restriction enzyme digestion as a result of gradual loss of restriction sites (Figure 4.2B) and by the shift in T_m during DSC analysis.

4.4.2 DSC Characterization of Nucleosomal DNA

When characterized by DSC under physiologically relevant salt condition, both S1 and 601 DNA show a single transition centered around 82 or 90 °C (Figure 4.3A), with an enthalpy change (ΔH) of around 1050 or 1400 kcal mol⁻¹, respectively. Changing the DNA concentration does not affect the T_m of the S1 DNA, although the intensity of the overall signal increases (Figure 4.3B). The 601 DNA is thermally and enthalpically more stable than the S1 DNA. This result is consistent with the base composition of each sequence. 601 DNA has a 54.5% GC content, which is almost 20% higher than the S1 DNA (35.6%); the higher GC content contributes to the higher T_m and larger enthalpy change. Besides the difference of thermal and thermodynamic stability, the two sequences also exhibit small difference in the shapes of their melting profiles. The melting curve of 601 DNA is more Gaussian in shape, while S1 DNA has a minor shoulder transition prior to the major transition. This difference is also reflected when we fit our data to a two-state model: the 601 DNA signal can be modeled by a single unfolding transition (Figure 4.3D), while the S1 DNA

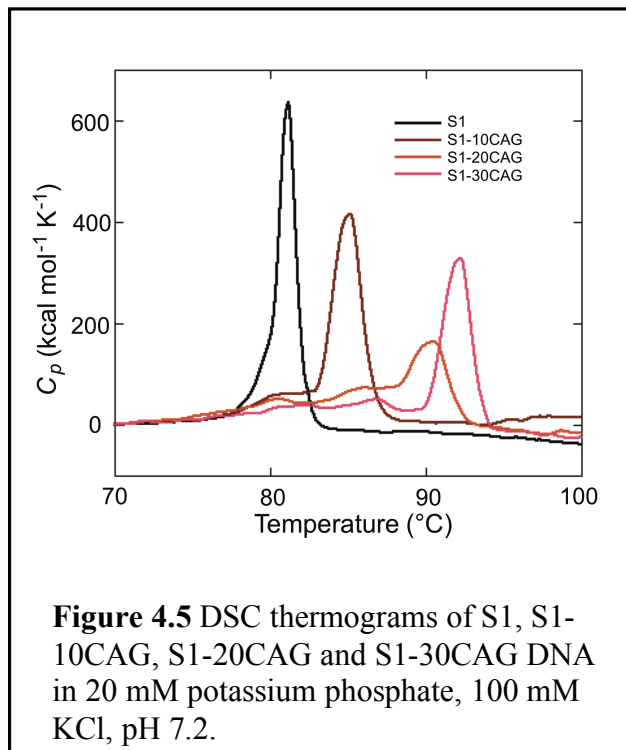
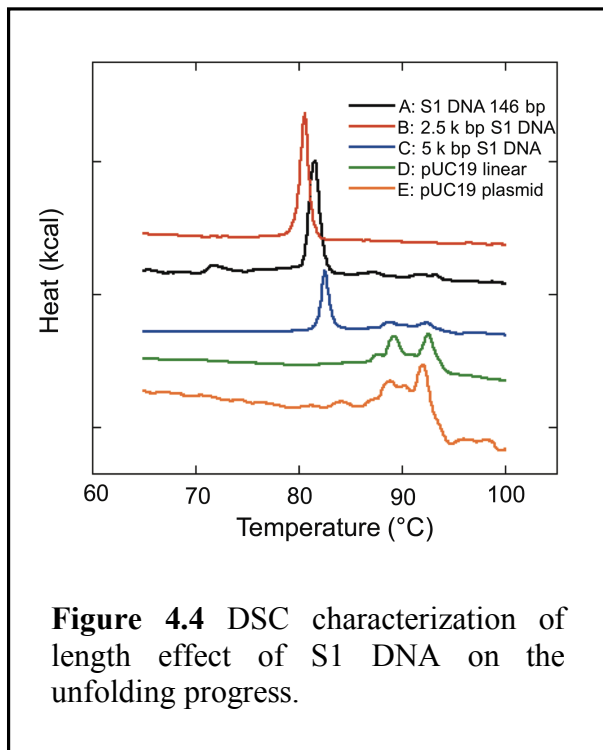
signal is best modeled by two unfolding transitions (Figure 4.3C). These results suggest that melting of 601 DNA is a more cooperative process and that S1 DNA is comprised of two distinct domains that melt at slightly different temperatures. It is generally thought that DNA unfolding



fails to follow the two-state unfolding model when the length is approaching a threshold.³³ In our case, S1 and 601 are almost identical in length. It is therefore possible to argue that the melting of DNA does not solely depend on the length of DNA, but also on whether the base composition is distributed evenly throughout the sequence. In order to support our argument, a 5 k bp plasmid

was prepared with 16 tandem S1 DNA repeats (2.5 k bp) inserted into the 2.5 k bp pUC19 backbone. The pUC19 plasmid itself, either circular or linear, showed multiple transitions from 88-95 °C (Figure 4.4). When the 16 tandem S1 DNA repeats (2.5 k bp) were inserted into the pUC19 plasmid, the pUC19 portion showed the same transitions as the unmodified pUC19 plasmid, although of much weaker intensity (Figure 4.4); however, a strong signal is also observed at around 82 °C, which is similar to the S1 DNA T_m (Figure 4.4). It is obvious that this transition corresponds to the melting of the 16 tandem S1 DNA repeats and that the synergic effect of all 16 S1 DNA domains contributes to the intense signal. The analyses of the plasmid DNA support our assumption that sequence distribution also affects the melting of DNA.

When replacing the central region of S1 DNA with 10, 20 and 30 of CAG repeats, the DNA becomes increasingly more stable overall compared with non-repeat S1 DNA, as indicated by the shift of T_m to higher temperatures (Figure 4.5). S1-10CAG DNA shows a major transition at 84 °C with a minor transition around 81 °C, which is the T_m of S1 DNA. S1-20CAG DNA showed a



major transition at 90 °C with several minor transitions from 80-88 °C, which correspond to the predominant transitions of S1 and S1-10CAG DNA. S1-30CAG shows a major transition centered at 92 °C. Surprisingly, the minor transitions are less prominent for the S1-30CAG DNA than for the S1-20CAG DNA. Differences in base distribution may also explain this observation. S1 DNA has an overall low GC content (35.6%). When replacing the central region with the high GC content CAG repeat (66.7%), it will introduce different sequence domains with T_m variations, and therefore several minor transitions occur in the S1 DNA containing different numbers of CAG repeats.

The multi-domain melting of DNA is also affected by the salt concentration. S1 DNA shows a well-defined melting curve and a corresponding T_m shift when increasing the salt concentration (Figure 4.6). A subsequent heating scan (2nd heating scan) following annealing can reproduce an identical melting curve under high salt conditions but not under low salt conditions. The existence of a shoulder peak in the 2nd heating scan under low salt conditions suggests that the S1 DNA has different sequence domains that melt at different T_m . The existence of multiple domains is probably due to an insufficient number of counter ions to fully stabilize duplex DNA, as suggested by the shape recovery at high salt conditions. The melting of 601 DNA also shows different shapes depending on salt concentration, but these shapes are not affected by the annealing process (Figure 4.6B). Two transitions are observed under low salt conditions, with a minor transition consistent at 90 °C and a major transition shift to higher T_m when increasing the salt. Starting from 100 mM KCl, the two transitions merge and only a single transition is observed (Figure 4.6B). We think that in the case of 601 DNA, the counter ions might selectively stabilize some sequence regions more than others, and divergence occurs. The above observation suggests that low salt conditions will result in less well-defined peak shape or even large peak splitting

depending on the DNA sequence. 100 mM salt seems to be the threshold to offer enough counter ion stabilization to DNA to restore the DNA melting curve to a single well-defined transition. As discussed previously for S1 containing different numbers of CAG repeats, we postulate that the existence of multiple transitions at low salt is due to the lack of counter ions. Increasing the salt concentration from 100 mM to 200 mM KCl in the S1 series does not merge the transitions into one, but every single transition moves towards higher temperature (Figure 4.6C). Therefore, the multiple transition observed in S1-10CAG, S1-20CAG and S1-30CAG are reflect sequence region heterogeneity rather than counter ion stabilization.

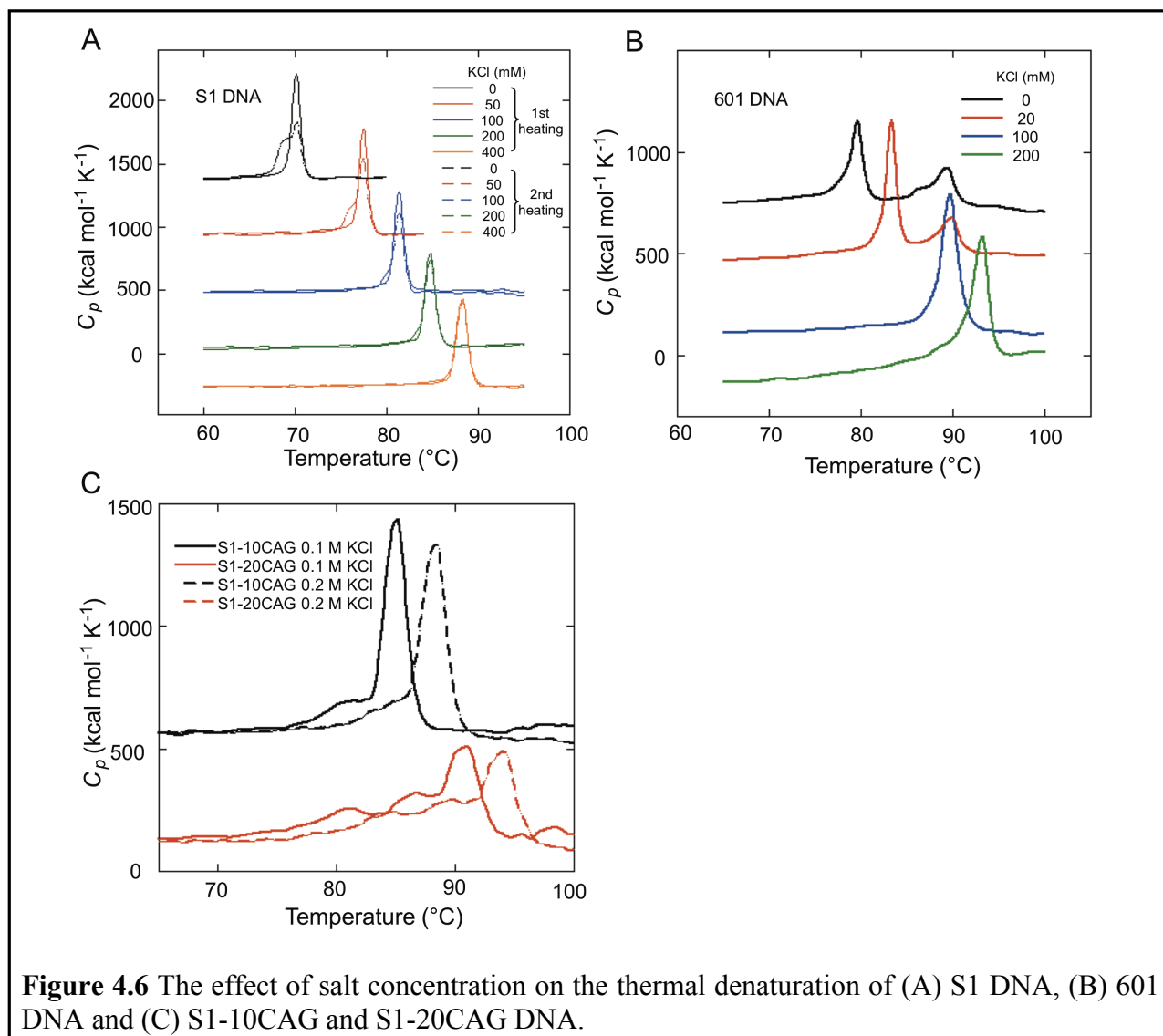
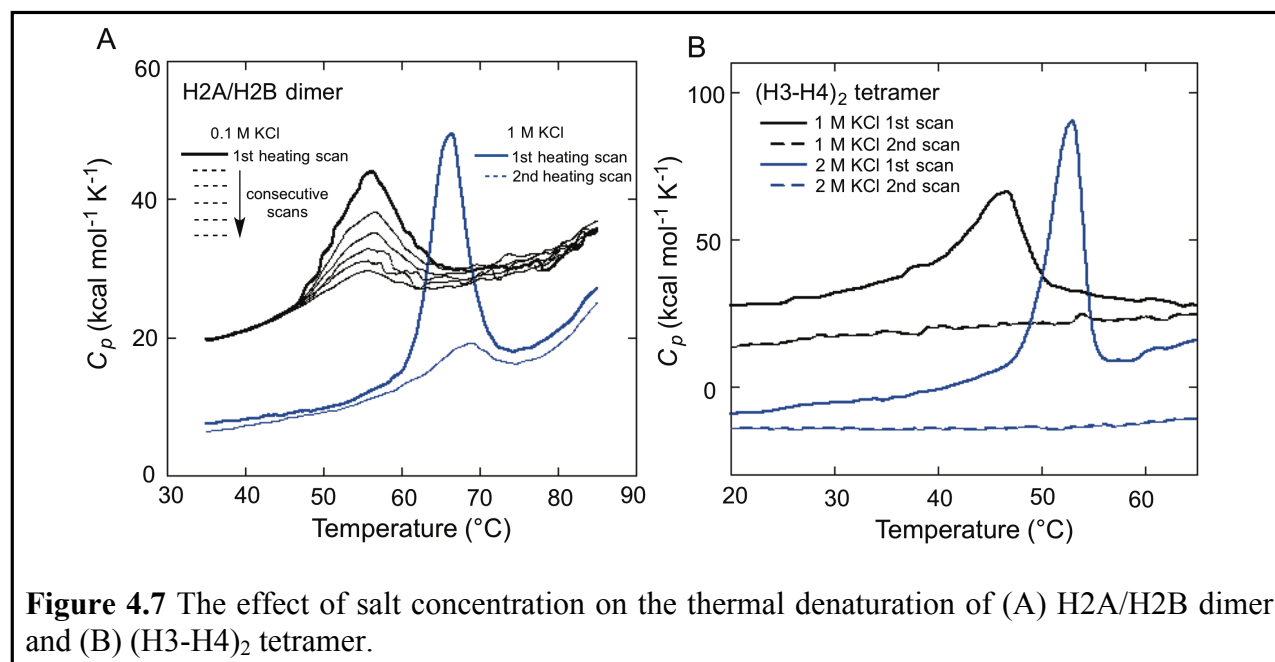


Figure 4.6 The effect of salt concentration on the thermal denaturation of (A) S1 DNA, (B) 601 DNA and (C) S1-10CAG and S1-20CAG DNA.

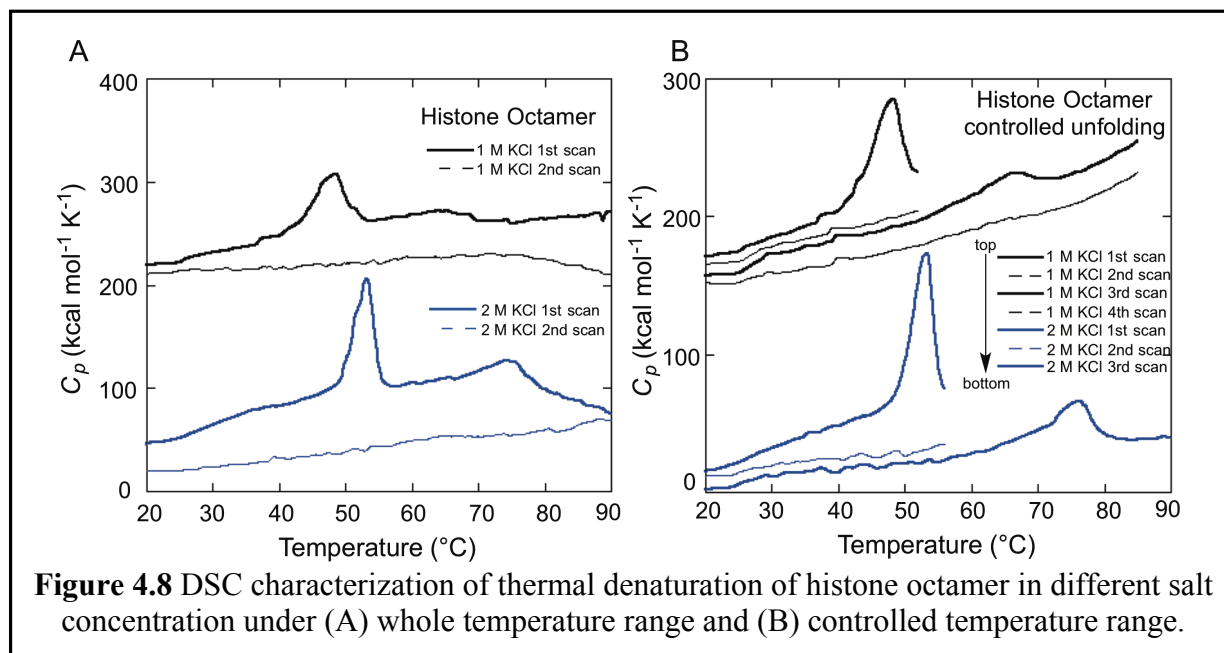
4.4.3 DSC Characterization of Histone H2A/H2B Dimer, (H3-H4)₂ Tetramer and Octamer

The histone octamer is composed of two copies of the H2A/H2B dimer and one copy of the (H3-H4)₂ tetramer. We therefore examined these two subunits. In 100 mM KCl, H2A/H2B shows a transition centered at around 56 °C. Subsequent scans demonstrate that unfolding of the H2A/H2B dimer is partially reversible, as indicated by the gradual loss of signal (Figure 4.7A). Increasing the salt concentration to 1 M KCl significantly enhances the thermal stability of the



H2A/H2B dimer, increasing the T_m to 68 °C. The increased salt prevents the H2A/H2B dimer from folding back to its native structure as indicated by the significant loss of signal during the second round heating scan. On the other hand, (H3-H4)₂ tetramer melts at much lower temperature compared with H2A/H2B dimer at the same buffer conditions. For example, in 1 M KCl buffer, (H3-H4)₂ tetramer has a T_m of 47 °C, more than 20 °C lower than the H2A/H2B dimer (Figure 4.7B). This result is consistent with previous reports that (H3-H4)₂ tetramer is thermally and thermodynamically less stable compared with the H2A/H2B dimer.^{19, 34} Increasing the salt to 2 M KCl shifts the T_m to 52 °C and gives a better defined transition. The thermal denaturation of (H3-

H4)₂ tetramer does not show any reversibility as subsequent heating scans do not generate any heat capacity change. The stabilities of the histone H2A/H2B dimer and the (H3-H4)₂ tetramer, isolated either from biological source or assembled from the purified recombinant histones, have been examined by several techniques, including spectrometric or DSC methods. Karantza and coworkers have characterized the pH and ion strength dependence of the thermodynamic stability of the H2A/H2B dimer by DSC,³⁵ and showed that the unfolding process is thermally reversible under low salt conditions within the neutral pH region. They also later investigated the (H3-H4)₂ tetramer and showed that the unfolding process is irreversible at any salt concentration, while its dimer form, H3/H4, is only reversible at low ionic strength and pH 4.5. Gloss and Placek were able to assemble the H2A/H2B dimer from recombinant proteins and showed that the H2A/H2B dimer also undergoes reversible unfolding when treated with urea.¹⁷ They concluded that the stability of the H2A/H2B dimer arises from screening the electrostatic repulsion of the highly charged histone tails.

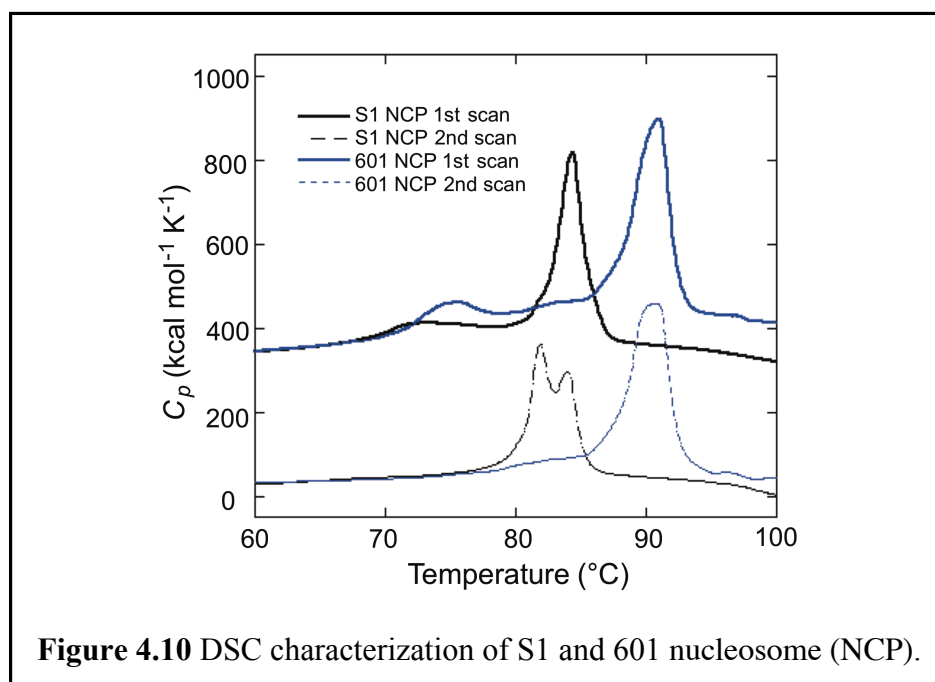
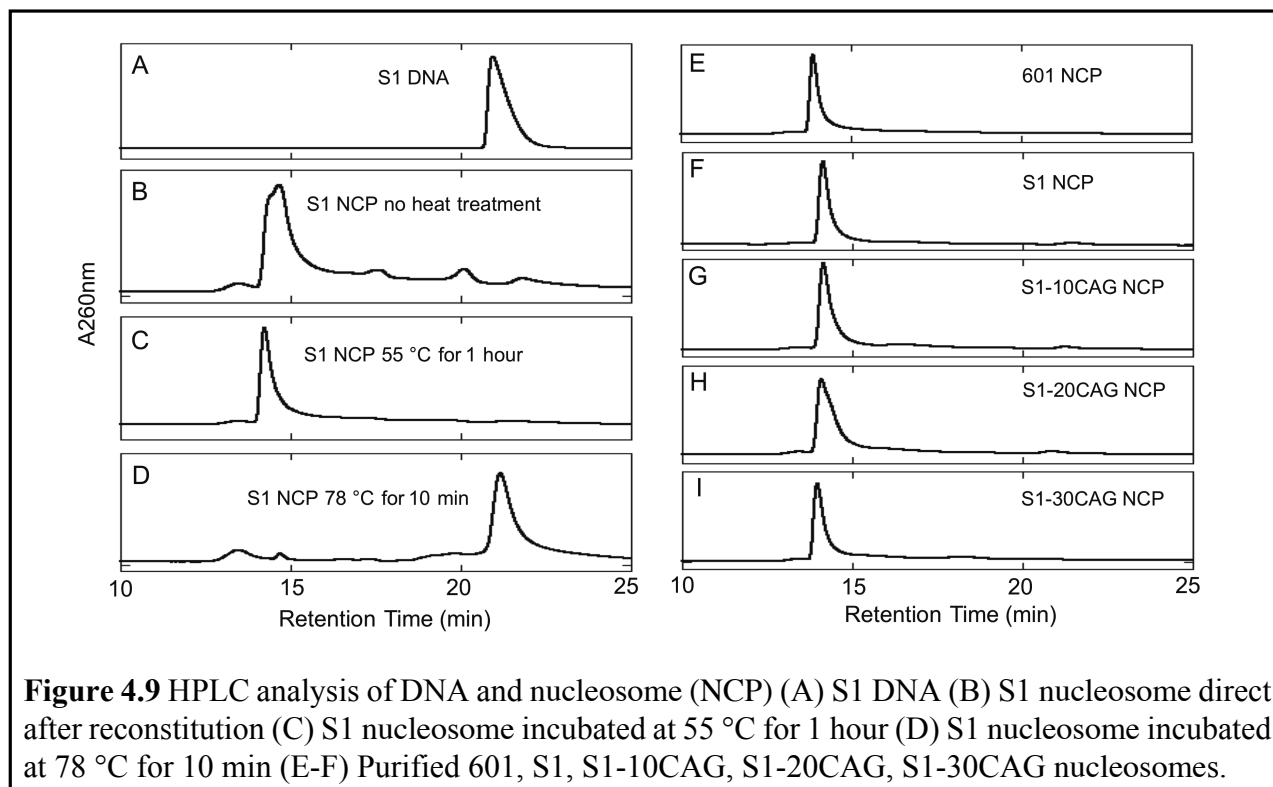


When the H2A/H2B dimer and the (H3-H4)₂ tetramer are assembled into the histone octamer, it is only stable under high salt conditions, and dissociates under low salt conditions. Therefore, we performed DSC on the histone octamer at 1 M and 2 M KCl (Figure 4.8A). For both salt conditions, we observed two transitions at 48 and 65 °C at 1 M KCl and 52 and 74 °C at 2 M KCl. These transitions are consistent with the results derived from the denaturation profiles of the individual subunits. The existence of two transitions suggests that histone octamer melts in stepwise fashion, with (H3-H4)₂ tetramer unfolding first and H2A/H2B dimer unfolding later. In order to confirm the stepwise melting mode, we performed a controlled-temperature DSC experiment on the histone octamer (Figure 4.8B). In this experiment, the scanning temperature was reversed immediately after the first transition (52 or 57 °C for 1 M or 2 M KCl condition), presumably the unfolding temperature of the (H3-H4)₂ tetramer. When the scan was repeated in the same temperature range, no signal was detected, suggesting that unfolding is irreversible. However, when the scan was performed across the whole temperature range (up to 90 °C), an additional transition appeared at 66 or 76 °C at 1 M or 2 M KCl, corresponding to H2A/H2B unfolding. Notably, the intensity of the unfolding transition of the (H3-H4)₂ tetramer is significantly larger than the unfolding transition of the H2A/H2B dimer. This difference in intensity is inconsistent with the results that the (H3-H4)₂ tetramer is thermodynamically less stable than the H2A/H2B dimer when these subunits are characterized individually.^{19,34} We think that the low temperature transitions observed during the histone octamer unfolding process do not only include the unfolding of the (H3-H4)₂ tetramer itself, but also the dissociation of the tetramer from the two H2A/H2B dimers. However, we cannot tell whether it is the unfolding of the tetramer that promotes the dissociation from the two dimers, or whether the tetramer and dimers dissociate first, followed by tetramer unfolding.

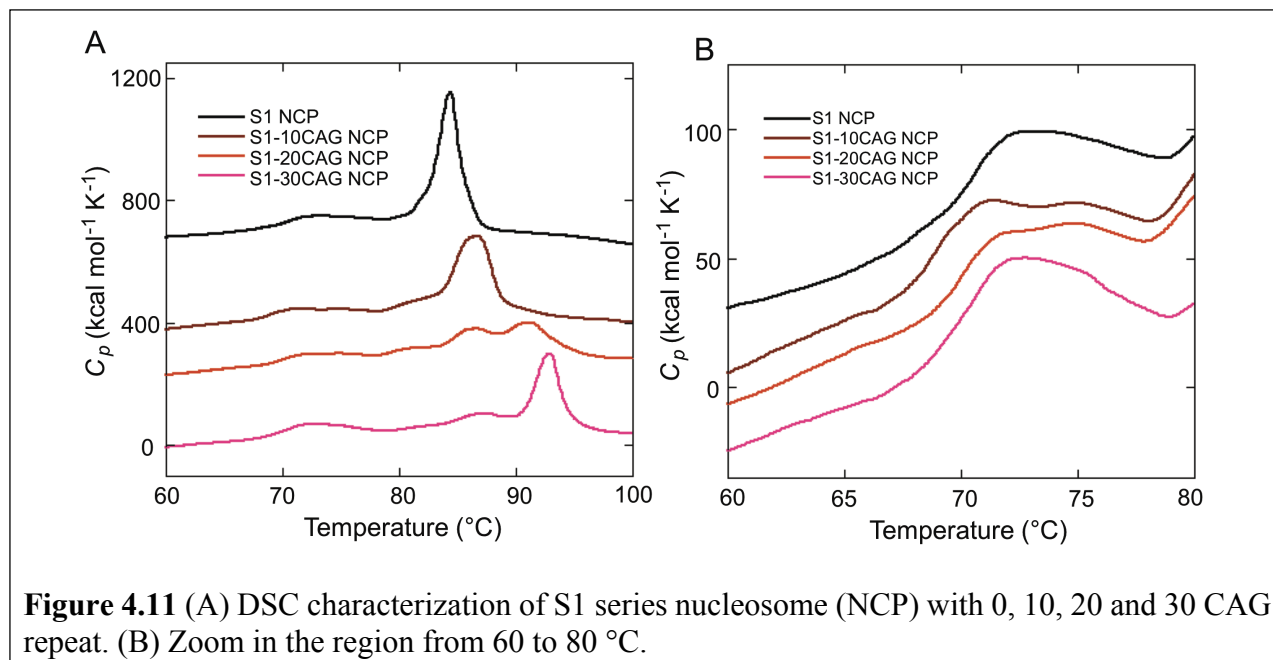
4.4.4 DSC Characterization of Nucleosome

With the prepared five different DNA sequences, we were able to reconstitute five different nucleosomes using the same procedure. HPLC with a DEAE column was used to separate the nucleosome from the unincorporated free DNA (Figure 4.9A). Immediately after reconstitution, some S1 series nucleosomes showed more than one population, with different translational positions as indicated by peak broadening or peak splitting in the HPLC chromatograms at 14 min (Figure 4.9B). After heat treatment at 55 °C for 1 hour, the S1 series nucleosomes converted to a homogeneous system as indicated by the sharp peak in HPLC (Figure 4.8C, 14 min). On the contrary, nucleosomes prepared with 601 DNA showed a single translational position, as indicated by the sharp peak in HPLC (Figure 4.9E, 14 min) without the heat treatment. With necessary heat treatment, homogenous nucleosomes with different DNA sequences were prepared for the DSC analysis (Figure 4.9E-I).

The DSC profile of the nucleosome with S1 DNA showed two transitions, with the minor transition centered at 72 °C and the major transition at 83 °C (Figure 4.10). With the identification of the major transition as the melting of S1 DNA compared with the free S1 DNA, we propose that the minor transition corresponds to the S1 DNA unwrapping from the histone core. In order to confirm our assignment, we also performed temperature-dependent HPLC analysis on the S1 nucleosome. S1 nucleosome was incubated at two temperatures either pre- or post- the assumed nucleosome unfolding transition for various times and then analyzed by HPLC (Figure 4.9C,D). At a temperature lower than the proposed nucleosome unfolding temperature, a single peak was observed that corresponds to the retention time of nucleosome (14 min); at temperature slightly above the proposed nucleosome unfolding temperature, the peaks disappeared at the nucleosome retention time (14 min), but another peak shows up that



to melting of free S1 DNA (21 min). The HPLC analysis confirms our assignment for the minor transition. Nucleosomes reconstituted with the strong positioning 601 DNA also show two transitions at 76 and 91 °C, respectively, which are 4 and 8 °C higher than the S1 nucleosome (Figure 4.10). Additionally, the unfolding of 601 DNA from the histone core has a better defined transition compared to the S1 nucleosome.



For nucleosome reconstituted with S1 series DNA with different numbers of CAG repeats, the results are slightly different from the 601 nucleosome but similar within the series (Figure 4.11). A weak lower temperature transition was observed at around 72 °C, followed by a strong characteristic DNA transitions similar to the free DNA with varying T_m depending on the CAG repeat number. Increasing the CAG repeat number within the S1 DNA does not seem to affect the unfolding of nucleosomes, since the 72 °C transition is similar in both T_m and intensity throughout the series (Figure 4.11B). Earlier studies on chromosome unfolding in cell nuclei or in whole cells by DSC showed conflicting results between different research groups, despite consistency in the number of thermal transitions and their corresponding T_m values.^{8, 9} It is possible that the

assignment of individual thermal transitions to chromatin conformational changes is challenging due to the complex nature of these systems. In summary, four transitions were observed in the thermal denaturation profile of nuclei or whole cell at 58-66 °C (transition I), 73-76 °C (transition II), 85-92 °C (transition III) and 105-107 °C (transition IV) under physiological salt conditions.^{6, 8-10, 14, 36, 37} Eligio and coworkers assigned transition I as the unfolding of scaffolding structure and proteins associated with heterogeneous nuclear RNA, transition II as the unfolding linker DNA region within the chromatin, transition III and IV as the unfolding of the nucleosome core particle placed within an expanded loop and the 30 nm fiber of chromatin, respectively,^{9, 10} their results included the extrapolation of previously published thermal denaturation data of nucleosomes or digested chromatin in low salt conditions (< 5 mM monovalent cations). Such analyses are possibly inaccurate due to the fact the DNA will exhibit very different thermal unfolding behavior with the protein components under low ionic conditions.³⁸⁻⁴⁰ On the contrary, various studies from Cole group gave a much more convincing explanation for assignment of these transitions. A combination of proteinase K, DNase I, linker histone, salt effects, and different cell types have been thoroughly examined to make the conclusion that transition II corresponds to the unfolding of the nucleosome, while transitions III and IV correspond to the melting of relaxed and supercoiled DNA.^{6, 8, 11, 12, 14, 15, 36, 37} Our results from the characterization of single nucleosomes, the basic unit of chromatin, support the assignment from Cole group. Although our system did not include linker histone H1, which is necessary for the higher order chromatin structure organization, it has been previously shown that the linker histone H1 does not contribute to the thermal stability of chromatin. When the H1 proteins are removed, there is no significant shifts of individual transitions.¹¹ It is surprising that our simple nucleosome system reflects the results of the isolated

chromatin, nuclei or the whole cells, considering that the biological samples are highly heterogeneous in terms of DNA sequence and histone protein post-modifications.

The DNA sequence-dependence of nucleosome formation has been examined extensively over the past decades.^{20, 21, 41-44} Interactions between DNA and the histone octamer includes the electrostatic interactions of the negative DNA phosphate backbone and the positively charged histone octamer core. The strength of the interaction depends on how closely these two components fit with each other, and therefore DNA sequences with the flexibility to bend are optimal for nucleosome formation. It is generally considered that there are two major sequence determinants affecting the bending property.⁴⁵ The bendable TA or AT dinucleotides, when occurring in a 10 bp periodic pattern, are in closer contact with the histones, allowing the DNA to fit closer to the histone core. While other stiff sequences, such as polyG/C or polyA/T, are often found in nucleosome depletion regions, which are important for promoter accessibility and transcriptional activity. Crystal structures and mechanistic studies on 601 nucleosomes reveal that the sequence preference and their bendable features facilitate DNA incorporation.^{46, 47} Although crystal structures are not available for our S1 series nucleosomes to give detailed structure preference of DNA incorporation, other methods have been developed to estimate the overall DNA sequence preference for nucleosome formation. Previous competitive nucleosome incorporation results from our lab showed that S1 DNA with more CAG repeats will facilitate the incorporation of DNA, with S1-30CAG showing the highest incorporation ratio.²⁴ The CAG repeat sequence is one of the naturally occurring sequences with high affinity towards the histone octamer, with a slight preference over the well-studied nucleosome positioning sequence *L. variegatus* 5S RNA gene.²¹ Therefore, we expect to see increased stability when more CAG repeats are introduced in the S1 sequence. Although we observed that the melting of DNA increases with increasing

numbers of CAG repeats, unfolding of the nucleosome does not seem to be affected by the number of CAG repeats. Competitive nucleosome incorporation studies on different nucleosomal DNA sequences showed that CAG repeat-containing DNA only exhibits a slight preference over the 5S RNA gene, with a free energy difference of around $0.53 \text{ kcal mol}^{-1}$ more stable. In contrast, the 601 nucleosomal sequence showed a much larger free energy difference when compared with 5S RNA, with a increase in stability of $2.89 \text{ kcal mol}^{-1}$. This is almost 1000-fold more favorable compared with the 5S RNA gene. It is probable that the DSC measurement is not sensitive enough, or that the DNA affinity to histone core needs to be significantly different, in order to observe a shift in the DSC thermogram for unfolding of the nucleosome.

4.5 Conclusion

In conclusion, we characterized the unfolding of histone H2A/H2B dimer, (H3-H4)₂ tetramer, histone octamer, nucleosomal DNA with different sequences and also their corresponding nucleosomes under various conditions by DSC. Understanding the unfolding process of the nucleosome and its subunits is the first step to understand the DNA dynamics in chromatin. It will also be interesting to see how other factors, such as histone post-translational modification or the presence of chromatin remodeling proteins, affect the stability of DNA in the nucleosome. Such experiments will provide a full analysis of nucleosome stability.

4.6 Reference

- (1) Luger, K., Mader, A. W., Richmond, R. K., Sargent, D. F., and Richmond, T. J. (1997) Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature* 389, 251-260.

- (2) McGhee, J. D., and Felsenfeld, G. (1980) Nucleosome structure. *Annu Rev Biochem* 49, 1115-1156.
- (3) Felsenfeld, G., and Groudine, M. (2003) Controlling the double helix. *Nature* 421, 448-453.
- (4) Cutter, A. R., and Hayes, J. J. (2015) A brief review of nucleosome structure. *FEBS Lett* 589, 2914-2922.
- (5) Andrews, A. J., and Luger, K. (2011) Nucleosome structure(s) and stability: variations on a theme. *Annu Rev Biophys* 40, 99-117.
- (6) Almagor, M., and Cole, R. D. (1989) In physiological salt conditions the core proteins of the nucleosomes in large chromatin fragments denature at 73 degrees C and the DNA unstacks at 85 degrees C. *J Biol Chem* 264, 6515-6519.
- (7) Nicolini, C., Trefiletti, V., Cavazza, B., Cuniberti, C., Patrone, E., Carlo, P., and Brambilla, G. (1983) Quaternary and quinternary structures of native chromatin DNA in liver nuclei: differential scanning calorimetry. *Science* 219, 176-178.
- (8) Touchette, N. A., and Cole, R. D. (1985) Differential scanning calorimetry of nuclei reveals the loss of major structural features in chromatin by brief nuclease treatment. *Proc Natl Acad Sci U S A* 82, 2642-2646.
- (9) Balbi, C., Abemoschi, M. L., Gogioso, L., Parodi, S., Barboro, P., Cavazza, B., and Patrone, E. (1989) Structural domains and conformational changes in nuclear chromatin: a quantitative thermodynamic approach by differential scanning calorimetry. *Biochemistry-U S* 28, 3220-3227.
- (10) Cavazza, B., Brizzolara, G., Lazzarini, G., Patrone, E., Piccardo, M., Barboro, P., Parodi, S., Pasini, A., and Balbi, C. (1991) Thermodynamics of condensation of nuclear chromatin.

- A differential scanning calorimetry study of the salt-dependent structural transitions. *Biochemistry-Us* 30, 9060-9072.
- (11) Touchette, N. A., and Cole, R. D. (1992) Effects of salt concentration and H1 histone removal on the differential scanning calorimetry of nuclei. *Biochemistry-Us* 31, 1842-1849.
 - (12) Ni, X., and Cole, R. D. (1994) Effects of Various Salts and pH on the Stability of the Nucleosome in Chromatin Fragments. *Biochemistry-Us* 33, 9276-9284.
 - (13) Russo, I., Barboro, P., Alberti, I., Parodi, S., Balbi, C., Allera, C., Lazzarini, G., and Patrone, E. (1995) Role of H1 in Chromatin Folding. A Thermodynamic Study of Chromatin Reconstitution by Differential Scanning Calorimetry. *Biochemistry-Us* 34, 301-311.
 - (14) Almagor, M., and Cole, R. D. (1989) Differential Scanning Calorimetry of Nuclei as a Test for the Effects of Anticancer Drugs on Human Chromatin. *Cancer Research* 49, 5561-5566.
 - (15) Almagor, M., and Cole, R. D. (1989) Changes in chromatin structure during the aging of cell cultures as revealed by differential scanning calorimetry. *Biochemistry-Us* 28, 5688-5693.
 - (16) McGinty, R. K., and Tan, S. (2015) Nucleosome structure and function. *Chem Rev* 115, 2255-2273.
 - (17) Gloss, L. M., and Placek, B. J. (2002) The Effect of Salts on the Stability of the H2A–H2B Histone Dimer. *Biochemistry-Us* 41, 14951-14959.
 - (18) Placek, B. J., and Gloss, L. M. (2002) The N-Terminal Tails of the H2A–H2B Histones Affect Dimer Structure and Stability†. *Biochemistry-Us* 41, 14960-14968.

- (19) Banks, D. D., and Gloss, L. M. (2003) Equilibrium folding of the core histones: the H3-H4 tetramer is less stable than the H2A-H2B dimer. *Biochemistry-Us* 42, 6827-6839.
- (20) Lowary, P. T., and Widom, J. (1998) New DNA sequence rules for high affinity binding to histone octamer and sequence-directed nucleosome positioning. *J Mol Biol* 276, 19-42.
- (21) Thastrom, A., Lowary, P. T., Widlund, H. R., Cao, H., Kubista, M., and Widom, J. (1999) Sequence motifs and free energies of selected natural and non-natural nucleosome positioning DNA sequences. *J Mol Biol* 288, 213-229.
- (22) Widom, J. (2002) Role of DNA sequence in nucleosome stability and dynamics. *Quarterly Reviews of Biophysics* 34.
- (23) Group, H. s. D. C. R. (1993) A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington's disease chromosomes. *Cell* 72, 971-983.
- (24) Volle, C. B., and Delaney, S. (2012) CAG/CTG repeats alter the affinity for the histone core and the positioning of DNA in the nucleosome. *Biochemistry-Us* 51, 9814-9825.
- (25) Simpson, R. T., Thoma, F., and Brubaker, J. M. (1985) Chromatin reconstituted from tandemly repeated cloned DNA fragments and core histones: A model system for study of higher order structure. *Cell* 42, 799-808.
- (26) Richmond, T. J., Searles, M. A., and Simpson, R. T. (1988) Crystals of a nucleosome core particle containing defined sequence DNA. *J Mol Biol* 199, 161-170.
- (27) Dyer, P. N., Edayathumangalam, R. S., White, C. L., Bao, Y., Chakravarthy, S., Muthurajan, U. M., and Luger, K. (2003) Reconstitution of Nucleosome Core Particles from Recombinant Histones and DNA. *Methods Enzymol* 375, 23-44.
- (28) Luger, K., Rechsteiner, T. J., and Richmond, T. J. (1999) Expression and purification of recombinant histones and nucleosome reconstitution. *Methods Mol Biol* 119, 1-16.

- (29) Studitsky, V. M., Clark, D. J., and Felsenfeld, G. (1994) A histone octamer can step around a transcribing polymerase without leaving the template. *Cell* 76, 371-382.
- (30) Ober, M., and Lippard, S. J. (2007) Cisplatin damage overrides the predefined rotational setting of positioned nucleosomes. *J Am Chem Soc* 129, 6278-6286.
- (31) Samadashwily, G. M., Raca, G., and Mirkin, S. M. (1997) Trinucleotide repeats affect DNA replication in vivo. *Nat Genet* 17, 298-304.
- (32) Pelletier, R., Krasilnikova, M. M., Samadashwily, G. M., Lahue, R., and Mirkin, S. M. (2003) Replication and expansion of trinucleotide repeats in yeast. *Mol Cell Biol* 23, 1349-1357.
- (33) Wu, P. (2000) Transition characteristics and thermodynamic analysis of DNA duplex formation: a quantitative consideration for the extent of duplex association. *Nucleic Acids Res* 28, 4762-4768.
- (34) Karantza, V., Freire, E., and Moudrianakis, E. N. (2001) Thermodynamic Studies of the Core Histones: Stability of the Octamer Subunits Is Not Altered by Removal of Their Terminal Domains†. *Biochemistry-Us* 40, 13114-13123.
- (35) Karantza, V., Baxevanis, A. D., Freire, E., and Moudrianakis, E. N. (1995) Thermodynamic studies of the core histones: Ionic strength and pH dependence of H2A-H2B dimer stability. *Biochemistry-Us* 34, 5988-5996.
- (36) Touchette, N. A., Anton, E., and Cole, R. D. (1986) A Higher-Order Chromatin Structure That Is Lost during Differentiation of Mouse Neuroblastoma-Cells. *J Biol Chem* 261, 2185-2188.
- (37) Almagor, M., and Cole, R. D. (1987) A high melting (105 degrees C) form of chromatin characterizes the potential of cells for mitosis. *J Biol Chem* 262, 15071-15075.

- (38) Weischet, W. O., Tatchell, K., Van Holde, K. E., and Klump, H. (1978) Thermal denaturation of nucleosomal core particles. *Nucleic Acids Res* 5, 139-160.
- (39) Bina, M., Sturtevant, J. M., and Stein, A. (1980) Stability of DNA in nucleosomes. *Proc Natl Acad Sci U S A* 77, 4044-4047.
- (40) Ausio, J., and Van Holde, K. E. (1986) Histone hyperacetylation: its effects on nucleosome conformation and stability. *Biochemistry-Us* 25, 1421-1428.
- (41) Widlund, H. R., Vitolo, J. M., Thiriet, C., and Hayes, J. J. (2000) DNA Sequence-Dependent Contributions of Core Histone Tails to Nucleosome Stability: Differential Effects of Acetylation and Proteolytic Tail Removal†. *Biochemistry-Us* 39, 3835-3841.
- (42) Gottesfeld, J. M., and Luger, K. (2001) Energetics and Affinity of the Histone Octamer for Defined DNA Sequences†. *Biochemistry-Us* 40, 10927-10933.
- (43) Tse, W. C., and Boger, D. L. (2004) Sequence-selective DNA recognition: natural products and nature's lessons. *Chem Biol* 11, 1607-1617.
- (44) Segal, E., and Widom, J. (2009) What controls nucleosome positions? *Trends Genet* 25, 335-343.
- (45) Struhl, K., and Segal, E. (2013) Determinants of nucleosome positioning. *Nat Struct Mol Biol* 20, 267-273.
- (46) Chua, E. Y., Vasudevan, D., Davey, G. E., Wu, B., and Davey, C. A. (2012) The mechanics behind DNA sequence-dependent properties of the nucleosome. *Nucleic Acids Res* 40, 6338-6352.
- (47) Vasudevan, D., Chua, E. Y., and Davey, C. A. (2010) Crystal structures of nucleosome core particles containing the '601' strong positioning sequence. *J Mol Biol* 403, 1-10.