PHOTO-INDUCED CHARGE FLOW THROUGH CANONICAL DNA/RNA AND NON-CANONICAL DNA CONTAINING TRINUCLEOTIDE REPEATS

by

Kaushiki Sen
M.Sc., Banaras Hindu University

THESIS
SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF SCIENCE

In the
Department of Molecular Biology and Biochemistry

© Kaushiki Sen 2010
SIMON FRASER UNIVERSITY
Spring 2010

All rights reserved. However, in accordance with the Copyright Act of Canada, this work may be reproduced, without authorization, under the conditions for Fair Dealing. Therefore, limited reproduction of this work for the purposes of private study, research, criticism, review and news reporting is likely to be in accordance with the law, particularly if cited appropriately.
Approval

Name: Kaushiki Sen
Degree: Master of Science
Title of Thesis: Photo-induced charge flow through canonical DNA/RNA and non-canonical DNA containing trinucleotide repeats.

Examining Committee:

Chair: Dr. Christopher Beh
Associate professor, Department of Molecular Biology and Biochemistry

Dr. Dipankar Sen
Senior Supervisor
Professor
Department of Molecular Biology and Biochemistry

Dr. Peter J. Unrau
Committee Member
Associate Professor
Department of Molecular Biology and Biochemistry

Dr. Melanie A. O’Neill
Committee Member
Assistant Professor
Department of Chemistry

Dr. Lisa Craig
Committee Member
Assistant Professor,
Department of Molecular Biology and Biochemistry

Dr. Jenifer Thewalt
Internal Examiner
Professor, Department of Molecular Biology and Biochemistry, Department of Physics

Date Defended/Approved: March 30, 2010
Declaration of Partial Copyright Licence

The author, whose copyright is declared on the title page of this work, has granted to Simon Fraser University the right to lend this thesis, project or extended essay to users of the Simon Fraser University Library, and to make partial or single copies only for such users or in response to a request from the library of any other university, or other educational institution, on its own behalf or for one of its users.

The author has further granted permission to Simon Fraser University to keep or make a digital copy for use in its circulating collection (currently available to the public at the “Institutional Repository” link of the SFU Library website <www.lib.sfu.ca> at: <http://ir.lib.sfu.ca/handle/1892/112>) and, without changing the content, to translate the thesis/project or extended essays, if technically possible, to any medium or format for the purpose of preservation of the digital work.

The author has further agreed that permission for multiple copying of this work for scholarly purposes may be granted by either the author or the Dean of Graduate Studies.

It is understood that copying or publication of this work for financial gain shall not be allowed without the author’s written permission.

Permission for public performance, or limited permission for private scholarly use, of any multimedia materials forming part of this work, may have been granted by the author. This information may be found on the separately catalogued multimedia material and in the signed Partial Copyright Licence.

While licensing SFU to permit the above uses, the author retains copyright in the thesis, project or extended essays, including the right to change the work for subsequent purposes, including editing and publishing the work in whole or in part, and licensing other parties, as the author may desire.

The original Partial Copyright Licence attesting to these terms, and signed by this author, may be found in the original bound copy of this work, retained in the Simon Fraser University Archive.

Simon Fraser University Library
Burnaby, BC, Canada
Abstract

The charge flow technique is used to study structure of nucleic acids. DNA-DNA duplexes and DNA-RNA heteroduplexes have been studied using charge transfer, yet little is known about RNA-RNA duplex conductivity. Both piperidine and aniline treatment on DNA-DNA duplexes with guanine-rich termini generated a large oxidative end-effect. Both DNA-DNA and RNA-RNA duplexes demonstrated similar findings. Sensitivity of the end-effect towards mismatches suggests its utility as sensitive reporter of mismatches.

Various diseases are associated with polymorphic DNA trinucleotide repeats (TNR). Disease associated TNRs (CAG, CGG, CCG, CTG) can form alternative DNA structures in vitro with two Watson-Crick base pairs sandwiching single mismatches. Mismatches interrupt base stacking and hence are known to affect duplex conductivity. Charge flow through duplexes carrying TNRs did not show any statistically significant difference in damage. The increase in damage from 10 °C to 21 °C, was not statistically significant. Our approach could not be applied to TNRs.
Acknowledgements

I would like to thank my supervisor, Dr. Dipankar Sen, for his relentless support, help and guidance throughout my graduate career at Simon Fraser University. I would never be able to express enough gratitude for everything that he has done for me. His belief in me and immense patience has helped me continue through the darker days.

I would also like to express my special gratitude to the members of my supervisory committee, Dr. Peter J. Unrau, Dr. Melanie A. O’Neill and Dr. Lisa Craig for all their important suggestions and encouragement throughout my degree. I shall always be grateful for the time that they have generously spent to support and guide me.

I would like to thank Dr. Jenifer Thewalt for taking time out of her busy schedule to be my internal examiner.

I wish to express my appreciation to Dr. Lucien Junior Bergeron for his guidance and help when I first joined the laboratory. His training during my first few months in the Sen laboratory has helped me to get a very good start with my first project. I owe a substantial portion of my research work to Dr. Bergeron.

I would like to thank all the past and present members of the Sen laboratory, particularly Dr. Edward Leung, Dr. Janet Huang and Nazanin Azrahimi for all the helpful suggestions. I also wish to thank Dr. Yong Liu, Gurpreet Sekhon, Lester Poon and Paul Cernak for their friendship and support.

I wish to thank the past and present members of the Unrau lab, for all the suggestion and help.

I would like to thank the staff of the department of Molecular Biology and Biochemistry for their help.

Finally, I wish to thank my family and friends for their patience, support and understanding.
# Table of Contents

Approval ............................................................................................................................... ii
Abstract ................................................................................................................................. iii
Acknowledgements .............................................................................................................. iv
Table of Contents .................................................................................................................... v
List of Figures ......................................................................................................................... viii
List of Tables ........................................................................................................................ x
List of Abbreviations ............................................................................................................ xii

Chapter 1:  General Introduction .......................................................................................... 1

1.1 Central Role of DNA ...................................................................................................... 1
1.2 Nucleic Acid Structures ............................................................................................... 3
  1.2.1 Base pairing ........................................................................................................... 3
  1.2.2 Helical Structure .................................................................................................... 6
  1.2.3 Higher Order Structures ....................................................................................... 8
1.3 Charge transfer ............................................................................................................. 11
  1.3.1 Monitoring Charge Transfer and Charge Flow Dependent DNA Cleavage through DNA ................................................................................................................. 11
  1.3.2 Various theories on the mechanism of charge transfer in duplex DNA ............... 20
  1.3.3 Charge transfer through DNA structures ............................................................. 23
  1.3.4 Application of charge transfer in sensing ............................................................. 24
  1.3.5 Biological applications of DNA mediated charge transport ............................... 25

Chapter 2:  Structural Studies of DNA/RNA by Photo-induced Charge Flow .................. 26

2.1 Introduction ................................................................................................................... 26
  2.1.1 The Potential of Utilizing Charge Flow and Quenching (CFQ) to Study DNA and RNA structures ................................................................. 28
2.2 Materials and Methods ............................................................................................... 31
  2.2.1 Synthesis of NHS ester of Anthraquinone-2-carboxylic Acid ................................ 31
  2.2.2 DNA Synthesis and Purification ......................................................................... 31
  2.2.3 C6-Amino DNA Synthesis and Purification ....................................................... 32
  2.2.4 Anthraquinone Coupling of C6-Amino DNA .................................................... 32
  2.2.5 HPLC Purification of Anthraquinone-coupled DNA .......................................... 33
  2.2.6 RNA synthesis ..................................................................................................... 35
  2.2.7 5’ End-Labeling .................................................................................................. 35
  2.2.8 Assembly of DNA or RNA Constructs ............................................................... 35
  2.2.9 Photo-irradiation of the DNA/RNA complexes ............................................... 36
  2.2.10 Iridium oxidation ............................................................................................... 36
2.2.11 G and A > G ladders ............................................................................ 36
2.2.12 Aniline and piperidine treatment .......................................................... 37
2.2.13 Denaturing Gel Electrophoresis ............................................................ 37
2.2.14 Data Analyses .......................................................................................... 38
2.3 Results and Discussion .............................................................................. 39
  2.3.1 Design of DNA/RNA duplex for charge transfer experiments .................. 39
  2.3.2 Comparison of piperidine and aniline cleavages of DNA ......................... 43
  2.3.3 To test that the products 1a, 1b and 1c are produced as a result of charge flow .............................................................................................................. 47
  2.3.4 Does the observed end-effect require a terminal G-C base pair? ............... 50
  2.3.5 The effect of 5' and 3'-end single-strand overhangs .................................. 54
  2.3.6 Is the end-effect sensitive to base pair mismatches within the double helix? 57
  2.3.7 Does an RNA double helix generate the same end-effect? ....................... 61
2.4 Conclusion .................................................................................................. 65
2.5 Future directions ......................................................................................... 66

Chapter 3: Structural Studies of DNA duplexes containing Trinucleotide Repeats by Photo-induced Charge Flow ........................................................................ 68

3.1 Introduction ................................................................................................. 68
  3.1.1 Structural information obtained from prior biophysical studies on trinucleotide repeat sequences ................................................................. 72
3.2 Materials and methods ............................................................................... 79
  3.2.1 Synthesis of NHS ester of Anthraquinone-2-carboxylic Acid .................. 79
  3.2.2 DNA Synthesis and Purification ............................................................... 79
  3.2.3 C6-Amino DNA Synthesis and Purification ............................................. 79
  3.2.4 Anthraquinone Coupling of C6-Amino DNA ......................................... 79
  3.2.5 HPLC Purification of Anthraquinone-coupled DNA ............................... 80
  3.2.6 5' End-Labelling ..................................................................................... 81
  3.2.7 Assembly of DNA Constructs .................................................................. 81
  3.2.8 Photo-irradiation of the DNA duplex assembly ....................................... 81
  3.2.9 Dimethyl Sulphate Modification of DNA (Guanine Ladder) ................. 82
  3.2.10 Piperidine Treatment and Denaturing Gel Electrophoresis ................ 82
  3.2.11 Data Analyses ...................................................................................... 82
3.3 Results and Discussion ............................................................................... 84
  3.3.1 Experimental Design of DNA duplexes containing trinucleotide repeats of the CNG type (N= Adenine/ Guanine/ Cytosine/ Thymine) ......................... 84
  3.3.2 Could DNA duplex carrying CCG, CGG, CTG and CAG repeats conduct charge? ........................................................................................................... 91
  3.3.3 How would any change in temperature during photo-irradiation affect the charge flow through DNA duplex? ............................................................... 96
3.4 Conclusion .................................................................................................. 107
3.5 Future Directions ...................................................................................... 108

Chapter 4: Final Conclusion ............................................................................ 109

Appendix: Common Experimental Procedures .............................................. 111
  Ethanol Precipitation ...................................................................................... 111
  5' End-labeling ............................................................................................... 111
  Dimethyl Sulfate Modification of DNA (G Ladder) .................................... 112
List of Figures

Figure 1.1: Watson-Crick base pairing................................................................. 2

Figure 1.2: RNA nucleotide versus DNA nucleotide........................................ 2

Figure 1.3: Hydrogen bonding sites of each base.................................................. 4

Figure 1.4: Non-Watson-Crick base pairs............................................................ 5

Figure 1.5: DNA triplet base pairing and a guanine quartet.................................... 6

Figure 1.6: Structures of double-stranded helices form by nucleic acids............... 8

Figure 1.7 Different structural motifs formed by DNA.......................................... 9

Figure 1.8 Guanine oxidation pathway.............................................................. 15

Figure 1.9 Alkali lability of modified nucleotides................................................. 16

Figure 1.10 Structure of some charge injectors used in studies of charge transfer ...... 19

Figure 1.11 Structure of anthraquinone (AQ) its derivatives charge injectors used for charge transfer ................................................................. 20

Figure 1.12 Mechanism of radical cation initiation and propagation during DNA charge transport................................................................. 22

Figure 1.13 Hole hopping over guanine and adenine............................................. 22
Figure 1.14  Polaron-like hopping model of charge transfer through DNA ........................................ 23

Figure 2.0  Mismatched G.T wobble base pairs ................................................................. 28

Figure 2.1  Anthraquinone coupled to DNA........................................................................ 33

Figure 2.2  (A) Schematic diagram of the model double helix, WT. (B) Sequences of the DNA WT duplex. (C) Anthraquinone coupled to the DNA.................................................. 42

Figure 2.3  DNA cleavage bands from aniline (A) and piperidine (B) workups of WT DNA, demonstrating Protocol 1 (W1) versus Protocol 2 (W2)................................................. 46

Figure 2.4  Control experiments to investigate whether the end-effect bands are charge-transfer generated, requiring AQ, light and aniline treatment .............. 49

Figure 2.5  (A) The end sequences of duplexes used in (B). (B) Denaturing 20% gel of the cleavage products from duplexes WT, D1, D2, and D3................................. 53

Figure 2.6  (A) The end sequences of duplexes used in (B). (B) Denaturing 15% gel of the cleavage products from duplexes WT and D4-D10.............................................. 56

Figure 2.7  Denaturing 20% gels showing the influence of mismatch substitutions on the end-effect bands............................................................................................... 60

Figure 2.8  Denaturing 20% gel comparing charge flow pattern through DNA and RNA WT duplexes. ................................................................................................................ 64

Figure 3.1  Constructs for DNA duplex carrying mismatched CNG triplet repeats, where the N:N mismatched pairs are G.G, A.A, T.T and C.C respectively......... 88

Figure 3.2  Constructs for Watson-Crick control duplexes carrying CNG repeats, where N = C, G, T, A ........................................................................................................... 90

Figure 3.3  15% Denaturing gel showing charge flow-dependent cleavage patterns in mismatched duplexes with corresponding Watson-Crick duplex controls.. ......... 95
Figure 3.4  Total cleavage for all four C.C, G.G, T.T and A.A mismatched duplexes and their W.C. controls.................................................................95

Figure 3.5  Construct for DNA duplex carrying mismatched CNG triplet repeats, where the N:N mismatched pair is T.T... .................................................................100

Figure 3.6  15% Denaturing gel showing cleavage patterns of CTG repeat containing Watson-Crick duplex control at different temperatures as a function of various photo-irradiation time periods.................................................................102

Figure 3.7  15% Denaturing gel showing cleavage patterns of CTG repeat containing T.T mismatched duplex at different temperatures as a function of various photo-irradiation time periods.................................................................104

Figure 3.8  Total cleavage for CTG mismatched and W.C control at 10, 21 oC for 30, 60, 120, 180 min .................................................................106
List of Tables

Table 1.1  Oxidation potentials of nucleosides ............................................. 13

Table 1.2  Calculated oxidation potentials for bases with reference to adjacent bases .......... 14

Table 3.1  Genomic loci affected by trinucleotide repeat instability. .............................. 69
# List of Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>8-Oxo-G</td>
<td>8-Oxoguanine</td>
</tr>
<tr>
<td>Ade (A)</td>
<td>Adenine</td>
</tr>
<tr>
<td>AQ</td>
<td>Anthraquinone</td>
</tr>
<tr>
<td>b.p.</td>
<td>Base pair</td>
</tr>
<tr>
<td>CFQ</td>
<td>Charge Flow Quenching Mapping</td>
</tr>
<tr>
<td>Cyt (C)</td>
<td>Cytosine</td>
</tr>
<tr>
<td>DCC</td>
<td>1,3-Dicyclohexylcarbodiimide</td>
</tr>
<tr>
<td>DEPC</td>
<td>Diethyl Pyrocarbonate</td>
</tr>
<tr>
<td>DMF</td>
<td>Dimethyl Formamide</td>
</tr>
<tr>
<td>DMS</td>
<td>Dimethyl Sulfate</td>
</tr>
<tr>
<td>DNA</td>
<td>Deoxyribose Nucleic Acid</td>
</tr>
<tr>
<td>dsDNA</td>
<td>Double Strand DNA</td>
</tr>
<tr>
<td>Dis/ dis</td>
<td>Distal</td>
</tr>
<tr>
<td>EDTA</td>
<td>Diaminoethanetetraacetic Acid</td>
</tr>
<tr>
<td>Gua (G)</td>
<td>Guanine</td>
</tr>
<tr>
<td>HPLC</td>
<td>High Pressure Liquid Chromatography</td>
</tr>
<tr>
<td>ISC</td>
<td>Intersystem Crossing</td>
</tr>
<tr>
<td>MgCl₂</td>
<td>Magnesium Chloride</td>
</tr>
<tr>
<td>MnCl₂</td>
<td>Manganese Chloride</td>
</tr>
<tr>
<td>Na₂B₄O₇</td>
<td>Sodium Borate</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
</tr>
<tr>
<td>NaCl</td>
<td>Sodium Chloride</td>
</tr>
<tr>
<td>NaOAc</td>
<td>Sodium Acetate</td>
</tr>
<tr>
<td>NH₂</td>
<td>Amino</td>
</tr>
<tr>
<td>NHE</td>
<td>Normal Hydrogen Electrode</td>
</tr>
<tr>
<td>NHS</td>
<td>N-hydroxysuccinimide</td>
</tr>
<tr>
<td>NMR</td>
<td>Nuclear Magnetic Resonance</td>
</tr>
<tr>
<td>PCR</td>
<td>Polymerase Chain Reaction</td>
</tr>
<tr>
<td>Pip</td>
<td>Piperidine</td>
</tr>
<tr>
<td>Prox/ prox</td>
<td>Proximal</td>
</tr>
<tr>
<td>RCF</td>
<td>Relative Centrifugal Force</td>
</tr>
<tr>
<td>RNA</td>
<td>Ribose Nucleic Acid</td>
</tr>
<tr>
<td>ss DNA</td>
<td>Single Strand DNA</td>
</tr>
<tr>
<td>TE</td>
<td>Tris-Cl, EDTA Buffer</td>
</tr>
<tr>
<td>TEAA</td>
<td>Triethylammonium acetate</td>
</tr>
<tr>
<td>Thy (T)</td>
<td>Thymine</td>
</tr>
<tr>
<td>TNR</td>
<td>Trinucleotide Repeat</td>
</tr>
<tr>
<td>Tris-Cl</td>
<td>2-Amino-2-(hydroxymethyl) propane-1,3 diol Chloride</td>
</tr>
<tr>
<td>Ura (U)</td>
<td>Uracil</td>
</tr>
<tr>
<td>UV</td>
<td>Ultraviolet</td>
</tr>
<tr>
<td>W.C.</td>
<td>Watson-Crick</td>
</tr>
</tbody>
</table>
Chapter 1: General Introduction

1.1 Central Role of DNA

One of the most important discoveries in the history of biological science has been the structure of the DNA double helix (Watson and Crick, 1953). The complementary base pairing of guanine to cytosine and adenine to thymine, shown in Figure 1.1, is the central basis of DNA replication, RNA transcription and protein translation. The rules of complementary base pairing is a characteristic of DNA that enables one strand of DNA to serve as a template for the synthesis of its complementary strand during DNA replication or for the synthesis of RNA during transcription. This ability to form complementary base pairs enables DNA to be the ideal heritable molecule for the storage of genetic information in living organisms. RNA and DNA are chemically similar molecules except that the thymine and deoxyribose sugar in DNA are being replaced by uracil and ribose sugar respectively in RNA (Figure 1.1.B and 1.2).
Figure 1.1: Watson-Crick base pairing. Each Watson-Crick base pair consists of a purine hydrogen bonding with a pyrimidine. A) Adenine forms two hydrogen bonds with thymine. B) In RNA, thymine is replaced by uracil and forms two hydrogen bonds with adenine. C) Guanine forms three hydrogen bonds with cytosine. The strand polarity is indicated by a (+) or (-), which represent the sugar phosphate backbone going into or coming out of the plane.

Figure 1.2: Comparison of DNA nucleotide against RNA nucleotide. A) The deoxyribose sugar in DNA has two hydrogen atoms at the C2’ position. B) The ribose sugar of RNA has a hydroxyl group and one hydrogen atom at the C2’ position. The presence of the hydroxyl group causes the ribose sugar to adopt a C3’-endo, C2’-exo sugar pucker while the deoxyribose sugar adopts a C2’-endo, C3’-exo sugar pucker.
1.2 Nucleic Acid Structures

1.2.1 Base pairing

The oxygen and nitrogen functional groups of each base act as hydrogen bond donors or acceptors. Figure 1.2 illustrates the potential. Nucleosides and nucleotides provide more hydrogen bonding opportunities through the 2’-hydroxyl of the ribose sugar and the phosphate group. The bases not only hydrogen bond with each other, but they can also hydrogen bond with polar amino acid residues, hydrated metal ions, and water molecules. Hydrogen bonds are one of the important chemical interactions found in nucleic acid and nucleoprotein complexes.

![Diagram of base pairing](image-url)
Figure 1.3: Hydrogen bonding sites for the four DNA bases. The functional groups of each base can act as hydrogen bond donors (shown by solid arrow pointing away from the base) or as hydrogen bond acceptors (shown by open arrow pointing towards the base).

Under physiological conditions, DNA is usually found in a double helix stabilized by Watson-Crick hydrogen bonds and base stacking interactions. The basic geometry of each Watson-Crick base pair is such that any sequence will not significantly distort the structure of the double helix. In this base-pairing scheme, guanine forms three hydrogen bonds with cytosine and adenine forms two hydrogen bonds with thymine in DNA or uracil in RNA (Figure 1.1). Aside from the Watson-Crick base pairs, the bases can form many other different types of base pairs. Figure 1.4 shows a sample of non-Watson-Crick base pairs involving at least two hydrogen bonds. These base pairs can be found in many different nucleic acid structures. For example, G•U base pairs are found in the tRNAPhe (Kim et al., 1974); G•U and A•C base pairs are found in the RNA component of RNase P (Chen and Pace 1997); and G•U, A•G, G•G, A•C, C•C, and U•U base pairs are found in ribosomal RNA (Gutell et al., 1993).

There are nucleic acid base pairings that involve more than two bases. Base triplets can be found in tRNAPhe (Kim et al., 1974), catalytic introns (Michel and Westhof, 1990), and DNA triplexes (Moser and Dervan, 1987). There is a unique tetrad called a guanine quartet that involves the Hoogsteen base pairing of four guanines. Guanine quartets are known to form from various guanine-rich DNA (Sen and Gilbert, 1988) and RNA (Kim et al., 1991) sequences. Figure 1.5 shows examples of base triplets and a guanine quartet.
Figure 1.4: Non-Watson-Crick base pairs. There are at least 27 different base pairs involving at least two hydrogen bonds (Saenger, 1984). Six such examples of non-Watson-Crick base pairs: G•U wobble; G•A N1-N1, carbonyl-amino (imion); A•C reverse hoogsteen; G•G N7-N1, carbonyl amino; C•C carbonyl-amino symmetric; and U•U 2-carbonyl-N3, 4-carbonyl-N3.
1.2.2 Helical Structure

The double helical model of DNA was first proposed in 1953 by Watson and Crick (Watson and Crick, 1953). Complementary bases from two DNA strands form hydrogen bonds in the interior of the helix and the sugar-phosphate backbone extends along the outside of the helix. The two strands run antiparallel with respect to each other. There are three main forms of nucleic acid helices: B-form, A-form, and Z-form (Figure 1.6). DNA can adopt all three forms of helices, but it is mostly found in the B-form. RNA
helices are predominantly found to be in the A-form. Not surprisingly, the sugar pucker plays an important role in determining the helical form in double-stranded DNA and RNA. The C3’-endo, C2’-exo sugar pucker of ribose favours the A-form helix and the C2’-endo, C3’-exo sugar pucker of deoxyribose favours the B-form helix. Both A-form and B-form are right handed helices formed from any nucleic acid sequence. Conversely, Z-form is a left-handed helix formed from alternating purine-pyrimidine sequences. The A-form helix is thicker and compressed along the helical axis while the Z-form is thin and elongated. The B-form helix falls in the intermediate region when compared to the two other forms. Other distinguishing features between all three helices are the grooves of each helix. In the B-form, the major groove is wide and the minor groove narrow, but both grooves are of similar depths (Bloomfield et al., 2000). The major and minor grooves of the A-form are similar in width, but the major groove is deep and minor groove is shallow (Bloomfield et al., 2000). Finally, the Z-form has a somewhat convexed major groove and a narrow but deep minor groove (Bloomfield et al., 2000).
Figure 1.6: Structures of double-stranded helices formed by nucleic acids. The A-form, B-form and Z form of helix. All the three different types of helices are formed by the same sequence.


1.2.3 Higher Order Structures

Nucleic acids can form many other elaborate structures besides the standard double helix (or duplex). Generally speaking, most of these diverse and complex structures have been associated with RNA. In the cell, long single stranded RNA molecules can fold into various complex structures to carry out a variety of different cellular functions, such as tRNAs, group I and group II self-splicing introns, and ribosomal RNAs. Single-stranded DNA molecules can also fold into a variety of similar and different structures when compared to ds DNA molecules. Figure 1.7 shows some simple structural motifs that DNA molecules can fold into. Single-stranded regions are nucleic acid sequences that are not involved in base pairing. In some cases, a single-
stranded region can be classified as part of a particular secondary structure motif such as a loop and a bulge.

Figure 1.7 Different structural motifs formed by DNA. The structures from D - G do not have to form from multiple strands. Any of the open ends can be closed by a single-strand DNA loop.

A simple secondary structural motif is the hairpin (Figure 1.7 C). This structure is formed when two complementary regions of a single molecule fold into a duplex. The
resulting structure is a helical region linked to a single stranded region. The single-stranded loop region may or may not influence the stability of the overall structure. The number of nucleotides in the loop can vary, but the shortest consists of three bases, such as that found in GNA-DNA tri-loops (Chou et al., 1996) and UGU RNA tri-loop (Flodell et al., 2002). The first and the last bases in each of these tri-loops are guanine and adenine, and uracil and uracil respectively. The base positioned in the middle in both the loops, ‘N’ could be any of the four bases. Other well-known loop sequences are the RNA “tetra loops” found in ribosomal RNA, most notably the highly stable “GNRA”, “UNCG” and “CUUG” tetra-loops (Woese et al., 1990).

Nucleic acid molecules can still form double helices when the two sequences are not perfectly complementary. The resulting structures are duplexes with mismatches or mismatch base pairs, internal loops, or bulges (Figure 1.7 D and E). Mismatched base pairs are non Watson-Crick base pairs as described previously. In order to minimize the frequency of deleterious effects caused by mismatched base pairs, cells have a variety of different enzymatic systems to correct these potentially serious mistakes (Mol et al., 1999; Li, 2008). The stability of the duplex may decrease to some degree depending to the type of mismatch. Mismatches are not restricted to a single base pair. Several continuous mismatches are typically referred to as an internal loop. There are two different types of internal loops: symmetric, where there are equal number of bases on both sides of the loop (Figure 1.7 E) and asymmetric, where one strand has more bases in the loop than the other strand. A variation of an internal loop is formed when the loop exists only on one strand. This is called a bulge. Bulges in DNA can be very useful because they allow duplex to form predictable kinks (reviewed in Lilley, 1995).
1.3 Charge transfer

1.3.1 Monitoring Charge Transfer and Charge Flow Dependent DNA Cleavage through DNA

Of the many interesting physical and chemical properties of DNA, charge transfer is one that has fascinated researchers for more than the past two decades. Upon ultraviolet irradiation, duplex DNA has been shown to facilitate charge transfer through its hydrogen-bonded base pairs to distances of 200 Å or more (reviewed in Joy and Schuster, 2005). Two classes of charge conduction have been studied so far: (1) a class in which the carrier of charge is a base radical cation (an “electron hole’’); and (2) a class in which it is a radical anion (an “excess electron”). Of the two, electron hole transfer has been studied to a greater extent.

Initial photochemical and photophysical studies were conducted using oligonucleotide assemblies containing electron donors and acceptors and the electron transfer was measured by fluorescence quenching as a function of distance (Núñez et al, 2002). Long distance charge transfer in DNA has been studied mainly through biochemical assays, time-resolved spectroscopic measurements, and electrochemical analysis.

The process of charge transfer in DNA has been related to various diseases like cancer, arteriosclerosis and aging (reviewed in Joy and Schuster, 2005). Hence, a clear understanding of the mechanism and factors governing charge transfer is of great importance. It has already been accepted that long-distance charge transfer in DNA is initiated by the photo-excitation of a covalently bound electron acceptor (photooxidant)
generating a radical cation that can travel throughout the duplex by a thermally activated hopping mechanism. The radical cations travelling across, damages the DNA mostly at positions containing guanine doublets and or guanine triplets (Schuster, 2000; Giese et al., 2001; Liu et al., 2004; O’Neill and Barton, 2004). The position of photooxidant attachment to the DNA was designed in such a manner, so that the photooxidant and guanine doublets and or triplets were spatially separated. This ensured proper monitoring of the charge flow through the duplex.

Any guanine base that transiently hosts the mobile radical cation is susceptible to a side reaction with water (and/ or dissolved oxygen), leading to the formation of guanine oxidation products such as 8-oxoguanine or diaminooxazolone (reviewed in Burrows and Muller, 1998). The formation of these products (Figure 1.8) can serve as a convenient biochemical method for monitoring the charge flow path through DNA, on treatment with hot piperidine (acidified aniline in the case of RNA). Piperidine or aniline breaks the DNA (or RNA) strand at the sites of such oxidative damage (Figure 1.9). The susceptibility of any particular guanine within a double helix to charge flow-related oxidative damage depends on the sequence context of the guanine in the helix (Núñez , 1999; Schuster 2000; Giese 2002).

Of the four DNA bases, guanine (G) is the most easily oxidized (Table 1.1) and the guanine radical cation, G•+, is the most stable of the four possible radical cations (Seidel et al. 1996). Hence, guanine would be the first base in a duplex from which an electron is being pulled out. The first G•+ thus formed would then pull out another electron from the next closest guanine in its vicinity and thus transfers the positive charge to the next guanine. The G•+ is the intermediate charge carrier in the hopping process and
charge trapping occurs at a stretch of adjacent Gs (for example 5’GG3’ or 5’GGG3’) (Sugiyama and Saito, 1996).

Stretches of adjacent guanines, such as 5’GG3’ or 5’GGG3’ are particularly susceptible to oxidation, with the 5’-most guanine of each stretch being most oxidizable (Table 1.2) (Burrows and Muller, 1998; Núñez, 1999; Schuster 2000; Giese 2002). Theoretical studies have shown that the guanine radical cation located at the 5’ side of a GG doublet would be highly stabilized by the adjacent guanine on the 3’ side of that doublet. The guanine at the 3’ end of a doublet cannot be stabilized by the guanine at the 5’ side.

Table 1.1 Oxidation potentials of nucleosides. The first five oxidation potentials were measured in aprotic solvent (acetonitrile) conditions (a - (Seidel et al., 1996). The remaining oxidative potentials were done in protic (aqueous) solvents at pH 7.0 (b - (Steenken and Jovanovic 1997) and pH 8.0 (c - (Yanagawa et al., 1992). The oxidation potentials of the bases are dependent on the solvent (aprotic versus protic) conditions, and therefore, hydrogen bonds will change the oxidation potentials of all nucleobases.

<table>
<thead>
<tr>
<th>Base</th>
<th>E°(V vs NHE)_a</th>
<th>E°(V vs NHE)_b</th>
<th>E°(V vs NHE)_c</th>
<th>pH</th>
</tr>
</thead>
<tbody>
<tr>
<td>Guanine</td>
<td>1.49</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adenine</td>
<td>1.96</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cytosine</td>
<td>2.14</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thymine</td>
<td>2.11</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Uracil</td>
<td>&gt;2.39</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rG</td>
<td>1.29</td>
<td>7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>rA</td>
<td>1.42</td>
<td>7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>dC</td>
<td>1.60</td>
<td>7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>dT</td>
<td>1.70</td>
<td>7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>r(8-oxoG)</td>
<td></td>
<td>0.58</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>r(8-oxoA)</td>
<td></td>
<td>0.92</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>r(5-OH-C)</td>
<td></td>
<td>0.62</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>r(5-OH-U)</td>
<td></td>
<td>0.64</td>
<td>8</td>
<td></td>
</tr>
</tbody>
</table>
Table 1.2 Calculated oxidation potentials for bases. The oxidation potential of the 5’ most guanine is influenced by its adjacent bases (a -Saito et al., 1995). The oxidation potential of 8-oxoguanine can also be influenced by its neighbouring bases (b- Prat, et al., 1998)

<table>
<thead>
<tr>
<th>Sequence</th>
<th>$E^\circ$(V vs NHE)$_a$</th>
<th>$E^\circ$(V vs NHE)$_b$</th>
</tr>
</thead>
<tbody>
<tr>
<td>GGG</td>
<td>0.64</td>
<td></td>
</tr>
<tr>
<td>GG</td>
<td>0.82</td>
<td></td>
</tr>
<tr>
<td>GA</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>GC</td>
<td>1.15</td>
<td></td>
</tr>
<tr>
<td>GT</td>
<td>1.16</td>
<td></td>
</tr>
<tr>
<td>G</td>
<td>1.20</td>
<td>0.85</td>
</tr>
<tr>
<td>8-oxoG</td>
<td></td>
<td>0.53</td>
</tr>
<tr>
<td>(8-oxoG)G</td>
<td></td>
<td>0.08</td>
</tr>
<tr>
<td>G(8-oxoG)</td>
<td></td>
<td>0.18</td>
</tr>
</tbody>
</table>

Once the strands are broken by piperidine or aniline, the fragmented products are separated from each other by denaturing electrophoresis. Densitometry calculation of the bands of the gel could analyse the damage intensity. Thus the phenomenon of oxidative DNA damage arising as a result of charge flow could be quantitatively mapped. Hence it was aptly called “charge flow dependent DNA cleavage” or “charge flow quenching mapping” (which would be referred to as CFQ henceforth). Thus charge transfer could prove to be a useful quantitative measure for mapping charge flow through double helices.
Figure 1.8  Guanine oxidation pathway. 8-oxoguanine is further oxidized to oxazolone, which is more piperidine labile when compared to 8-oxoguanine.

Figure 1.9  Alkali lability of modified nucleotides. Deprotonated deoxyguanine reacting with hot piperidine results in the β- and δ-elimination.


On investigating, the final products of the water reaction were found to be imidazolone and oxazolone that are far more alkali labile than 8-oxoguanine by itself. Though 8-oxoguanine is alkali labile, it is so only to a certain extent. The guanine oxidation product of 8-oxoguanine can never be completely alkali labile. In order to make the oxidation products more alkali labile, they have to be further oxidised by treating with an iridium (IV) compound. This is followed by the alkali treatment with hot aniline/piperidine is carried out.
It is possible to interpret the level of CFQ at any DNA site as the product of two distinct processes: (1) the efficiency of charge transfer to and from the guanine at that site (the rate of charge transfer) versus (2) the efficiency of the water reaction with the G•+ species generated at that site (the rate of the water oxidation reaction). In the case of a DNA double helix, the rate of charge transfer, step (1), is faster than the rate of the water oxidation reaction, step (2). The rate of charge/ hole transfer from a donor to an acceptor over long distances is in the range of $10^7$ - $10^{11}$ sec$^{-1}$, while over shorter distances, it is in a higher range of $10^9$ – $10^{12}$ sec$^{-1}$. The rate of reaction of G•+ with water can range only between milliseconds and microseconds in a duplex (reviewed in Giese, 2002; Wagenknecht, 2005). Its only in the cases of unusual geometries of nucleic acid molecules, can the water reaction be faster than the hole transfer.

This provides an explanation for the propagation of the charge through the DNA double helix. The level of observed damage at a particular base within a complexly folded nucleic acid reflects a balance between the rate of charge flow into, and out of, that base relative to the rate of water reaction of the resulting radical cation. The rate of intrastrand charge transfer in B-DNA duplexes is approximately 100 times faster than the rate of interstrand transfer (Kelley and Barton, 1999). Intrastrand base stacking occurs preferentially in B-DNA as compared to interstrand stacking; hence intrastrand charge transfer is faster than the other.

It should be possible to distinguish between highly oxidized guanines that are located within helices from those that are extrahelical and more significantly exposed to the solvent and it was speculated that enriching the solvent with reducing agents might help to quench the G•+, species and, in turn, prevent the water reaction and oxidative
damage. Charge flow dependent oxidative damage of the bases might also be able to provide some structural information regarding the degree of solvent exposure when compared to intrahelical bases as a standard.

In the last decade, the mechanism of DNA oxidation was being studied using a number of different compounds that were either intercalated between the base pairs or bound to a groove of the DNA or would have been capped on the terminus of the helix. Barton’s group has used rhodium and ruthenium intercalators to study charge transfer (1 and 2 respectively in Figure 1.10 below) (Núñez and Barton, 2000). Lewis and his group have employed stilbene linked hairpins to perform their charge transfer experiments (3 in Figure 1.10). Trioxatriangulenium ion (TOTA+) also has been used for the oxidation of DNA (Chen and Pace, 1997).
Schuster’s group has used various anthraquinone derivatives (Figure 1.11) covalently attached to the 5’ end of the DNA duplex to initiate charge transfer. The anthraquinone derivatives are end-capped, stacked onto the duplex which allows electric contact with the \( \pi \)–electrons of the DNA without causing any structural distortion to the DNA (reviewed in Joy and Schuster, 2005). Our laboratory uses anthraquinone which is covalently attached to the 5’–end of the DNA. The 5’ end of the DNA is modified with a
C6 carbon amino linker. The amino group attached to the C6 carbon linker is covalently attached to anthraquinone by a peptide bond (Fig 2.1 of section 2.2.4).

![Figure 1.11](image-url) Structure of anthraquinone (AQ) and its derivatives charge injectors used for charge transfer. (Joy A. and Schuster GB. 2005. Long-range radical cation migration in DNA: Investigation of the mechanism. *Chem. Commun.* **14**(22): 2778-2784).

### 1.3.2 Various theories on the mechanism of charge transfer in duplex DNA

There are two basic mechanisms that have been well established in the field of light-induced charge transfer in DNA: 1) the super-exchange mechanism; and 2) the hopping model (reviewed in Wagenknecht, 2006). Other alternative mechanisms also had been proposed earlier. The molecular wire model was one such mechanism where both the donor and acceptor molecules were organometallic intercalators. The charge is injected by the donor to the multi-base DNA bridge. The charge of the radical cation is localized over this multi-base DNA bridge and moves coherently in a rapid single step to the acceptor molecule. Experiment results based on fluorescence quenching of organometallic intercalators could only be explained by this mechanism (reviewed in Joy and Schuster, 2005).
The super-exchange mechanism is expected to operate only over a shorter distance scale, ranging up to a maximum of four A-T base pairs. The rate of charge transfer in super-exchange mechanism is completely distance dependent. Here the charge tunnels in one coherent step from the donor to the acceptor, but never resides on the DNA bridge between them (reviewed in Wagenknecht, 2006).

The hole hopping mechanism is characterized by multi-step hopping of the charge between the guanine bases. The individual hops between the sequential guanines take place by the super-exchange mechanism. The radical cation tunnels across the A-T base pairs separating these guanines but the charge is never localized on the A/T bridge. (Figure 1.13); (reviewed in Joy and Schuster, 2005). The hopping mechanism explains long-distance charge transfer up to a distance of 200 Å (Henderson, 1999; Núñez, 1999; reviewed in Joy and Schuster, 2005). It is outlined by the three steps of charge injection, charge transport, and charge trapping (Figure 1.12).
In the figure, AQ is the photo-oxidant used for initiating charge transfer. ISC is a radiationless process involved in a transition between two electronic states with different spin multiplicity.

In addition to multi-step hole hopping, a “polaron-like” or “phonon-assisted polaron-like” model of hole hopping was suggested by Schuster and his research group (Fig. 1.14). This model was based on the assumption that oxidation of a base within a DNA sequence would be accompanied by a structural distortion that results in lowering of the system’s energy. A polaron is a structural distortion within the DNA which
stabilizes and delocalizes the charge of the G•+ over several DNA bases. Formation of G•+ is accompanied by a subsequent structural change that lowers the energy for the system. This creates a shallow energy minimum where the G•+ gets trapped. Thermal motion of the duplex and its environment activates the hopping of such polarons (Henderson, 1999; Schuster, 2005, reviewed in Wagenknecht, 2006).

Barton and her group describe charge transfer as a hopping process between conformationally gated domains of well-stacked base pairs (O’Neill and Barton, 2004; reviewed in Wagenknecht, 2006). A domain is a transiently extended π–orbital over which the charge can delocalize. The charge delocalizes over a domain without getting trapped or making any structural change to the domain. Such delocalized domains may transiently form and break, and facilitate or inhibit charge transfer (O’Neill and Barton, 2004).

![Figure 1.14](image)

Figure 1.14  Polaron-like hopping model of charge transfer through DNA.


### 1.3.3 Charge transfer through DNA structures

It is expected that the integrity of π-stacking between adjacent bases or base pairs would strongly influence the efficiency of hole transfer in DNA. Where a π-stack is perturbed (such as with mismatches or bulges), hole migration efficiency decreases
noticeably (Barton 1999; Barton 2001). Therefore, the technique of charge transfer and quenching mapping provides a simple methodology to map charge transfer through intact DNA helices and as well as simple assemblages of helices, such as DNA three-way and four-way helical junctions (Odom, 2001; Sen 2002; Schuster 2003).

1.3.4 Application of charge transfer in sensing

The sensitivity of DNA charge transfer towards base stacking forms the basis for sensor applications. This property could be exploited to detect single base mismatches or base lesions by performing electrochemistry experiments on DNA films. In these assays, DNA duplexes were attached to a gold surface by a thiol linker. The DNA duplexes were also bound to a redox-active intercalator. The charge transport through such DNA-modified electrode surfaces was amplified using electrocatalysis. Monitoring the reduction of the intercalator bound to the DNA film, could then detect the presence of any single base mismatched pair. Such assays have a higher signal: noise ratio which increases the efficiency of mismatch discrimination. It is completely a new technology for the rapid detection of single nucleotide polymorphisms (Kelley et al., 1997; Kelley et al., 1998; Sam et al., 2001; Kelley et al., 1999; Kelley et al., 1999; Boon et al., 2000).

The electrochemical reduction of DNA-bound intercalators has also been shown to be useful in developing assays for examining protein-induced changes in DNA structure. Electrochemical reduction of a redox-active DNA intercalator gets diminished upon the binding of proteins that disrupt base pair stacking within the DNA duplex. Thus DNA mediated charge transfer could act as a novel probe for protein–DNA interactions and can be used in screening for inhibitors of protein–DNA interaction. It might also be
able to monitor real-time perturbations in DNA structure that are associated with protein binding and reaction (Boon et al., 2001).

1.3.5 Biological applications of DNA mediated charge transport

Implications of DNA mediated charge transport inside the cell has been shown by preferential base damage at the 5’-guanine of 5’-GG-3’ sites, a typical indicator of damage by DNA-mediated charge transfer (Núñez et al., 2001). DNA within the cell nuclei have been probed for charge transfer by using photoactivated rhodium intercalators. The size of the DNA packaged inside a nucleosome within the cell, is in the same range as that of the length of DNA up to which charge migration can occur. In nucleosome core particles, oxidative damage has also been demonstrated at a distance from the site of the photo-oxidant (Núñez and Barton, 2002). Moreover, for transcriptionally active DNA inside the nucleus, oxidative damage was found at protein-bound sites that were inaccessible to electron acceptors, as shown by photofootprinting. Electrochemistry experiments on DNA charge transfer have been successful in detecting mismatches and lesions within the DNA (Rajski et al, 2000). Hence in the future, one could imagine DNA charge transport as a sensing mechanism for detecting mismatches and lesions in DNA within the cell (Núñez et al, 2001). Researchers from Barton’s group have shown that repair of thymine dimers within DNA might be triggered oxidatively from a distance (Dandliker et al., 1997; Dandliker et al., 1998; Vicic et al., 2000). Moreover, thymine dimer repair has also been demonstrated reductively with bound flavins (Carell and Epple, 1998, Carell, 1995). Scientists have also proposed that cellular regions such as CpG islands and telomeres that have highly disproportionate guanine content might be representing hot spots for DNA damage (Freidman and Heller, 2001).
Chapter 2:  
*Structural Studies of DNA/RNA by Photo-induced Charge Flow*

2.1 Introduction

DNA mediated charge transfer is extremely sensitive to π-stacking of the bases in DNA. The significance of stacking of the electron acceptor and nucleotides within the DNA base stack has been highlighted in studies involving photo-induced charge flow between the modified bases in DNA (Núñez et al., 1998). DNA mediated charge flow is highly dependent on the stacking of the base pairs in between the points of initiation and termination of charge migration. Hence, charge transport studies might be able to provide a measure of the sequence dependent conformational dynamics and local flexibility of DNA.

Base bulges, flexible sequences and protein-induced distortions can greatly influence the efficiency of charge flow (Núñez et al., 1999; Hall and Barton, 1997). This sensitivity towards mismatches is the basis for discrimination of mismatches from hydrogen-bonded complementary Watson-Crick base pairs. Mismatches are usually stacked in DNA but are capable of undergoing greater dynamic motions than well-paired bases. DNA charge transfer detection which is dependent on the electronic coupling within the base stack is sensitive to such motions (Hunter et al., 1998; Hall and Barton, 1997).
Single-crystal X-ray diffraction studies on oligomers containing G.T mismatched base pairs show that they crystallize in the B-DNA-type conformation. G.T pairs are able to form stable base pairs linked by two hydrogen bonds. The G.T pairs can adopt a “wobble” structure with the thymine projecting in the major groove and guanine into the minor groove. The sugar phosphate backbone of DNA containing G.T pairs is sufficiently flexible to accommodate these mismatched pairs by small local adjustments in its conformation (Hunter et al., 1987). Although the stacking of bases is different from base stacking in Watson-Crick duplexes, it is quite extensive between G.T mispairs and adjacent base pairs, and is thought to stabilize such mismatch-containing DNA duplexes (Hunter et al., 1987).

Stacking of purines on pyrimidine bases exhibit better degree of overlap than that of pyrimidines stacked over purine bases (Hunter et al., 1987). Changes in the stacking behaviour are highly localized at the mismatch sites only. G.T mismatched pairs have similar widths of major and minor groove distances when compared to their native Watson-Crick base pairs (Fig 2.0). They have the same geometry in A, B and Z conformation of DNA. Hydration of DNA containing G.T pairs is similar to their corresponding Watson-Crick native DNA duplexes and could be the reason for the stabilization of G.T pair containing duplexes. The hydrogen bonding between the free functional groups of mismatched pairs and solvents is different from those formed by Watson-Crick base pairs (Hunter et al., 1987; Peyret et al., 1999).
Fig 2.0 Mismatched G.T. wobble-base pair showing structural similarity to Watson-Crick G-C and A-T base pairs.


Parallel investigations have attempted to unravel aspects of conformational transitions, helix dynamics, and ligand-DNA interactions by high-resolution nuclear magnetic resonance (NMR) studies in solution (Allawahi and Santa Lucia, 1998; Hunter et al., 1986).

2.1.1 The Potential of Utilizing Charge Flow and Quenching (CFQ) to Study DNA and RNA structures

The ability of DNA duplexes to conduct charge flow through base mismatches has been systematically examined in DNA assemblies using guanine oxidation ratios as a
measurement of charge transport efficiency (Núñez et al., 1999; reviewed in Boon and Barton, 2002).

Benefits of choosing CFQ as our structural probing tool are outlined as below:

1. CFQ – is a simple gel-based experimental technique that can provide structural information about nucleic acids (such as co-axial helical stacking conformations, changes in helical stacking conformation, and the degree of solvent exposure of the nucleobases) which are not easily obtained from other biophysical studies of X-ray crystallography, NMR, chemical and enzymatic probing, fluorescence quenching to name a few.

2. This is a simple gel-based new approach that could be applied to study the structures of nucleic acids and re-confirmed with the results from NMR, X-ray crystallography.

Helical stacking helps in modelling of complex structures such as in enzyme-substrate complexes. Change in solution conditions influences the helical stacking and such complexes could have alternate folding units (Leung and Sen, 2007).

Structural studies of RNA duplexes are thought to have a bigger application towards therapeutics. RNA Transcription amplifies the number of copies of any gene; hence the functional significance of RNA is greater than that of DNA. Unravelling secondary structures of RNA might open a new path towards therapeutics. CFQ studies have been systematically done on a variety of DNA double helices. Hence, a lot of information about their conductivity pattern and their implications on elucidating structures is already available to us. However, to our knowledge, CFQ has never been used to study structures of purely helical duplex/ complex RNA assemblies. Hence the
question could be put forward that whether CFQ also be able to give us more information about the secondary structures of RNA helices?

RNA structural probing has earlier been carried out by Schuster’s and Barton’s groups. Studies on RNA-DNA heteroduplexes and chimeras have been shown to conduct charges. Schuster’s group demonstrated charge flow through DNA/RNA chimera hybrid duplexes (Sartor et al., 1999). Chimera duplexes are duplexes in which both RNA and DNA nucleotides are present in the same oligonucleotide strand. The complementary AQ-coupled strand used in their charge transfer experiments was an oligonucleotide made of DNA nucleotides only; hence they called hybrid chimera duplexes. Barton’s group further advanced the field by demonstrating charge flow through duplexes made of a chimera strand containing a mixture of both DNA and RNA nucleotides and a complementary strand which having DNA nucleotides only. Moreover, they had also tested DNA/RNA hybrid duplexes that were assembled with one strand of DNA and a complementary strand of RNA (Odom and Barton, 2001). So, it would not be incorrect to say that CFQ shows potential as an experimental technique for DNA and RNA structural probing (Odom and Barton 2001; Sartor et al., 1999; O’Neill and Barton 2002).

The most important requirement for CFQ is the presence of helical stacking of bases in a nucleic acid molecule. This implies that theoretically both B-DNA and A-RNA should be able to conduct charges. This is the basic hypothesis for all our experiments. My aim is to study the relative conductive properties of DNA-DNA and RNA-RNA helical duplexes using CFQ as my tool.
2.2 Materials and Methods

2.2.1 Synthesis of NHS ester of Anthraquinone-2-carboxylic Acid

To label DNA with 5’ amino modification with anthraquinone, a derivative of anthraquinone with carboxylic acid, (namely anthraquinone-2-carboxylic acid) was directly coupled to a 5’ amino group on the end of C6 linker of the DNA. The coupling was done by the formation of a peptide bond, which ultimately led to the formation of anthraquinone labelled DNA (Fahlman and Sen, 2001).

2.2.2 DNA Synthesis and Purification

The DNA oligonucleotides were synthesized at the University of Calgary Core DNA Services. The oligonucleotides were treated with 10% piperidine at 90 °C for 30 min followed by lyophilisation. Such a pre-treatment removes base cleaved DNA molecules damaged during synthesis. This leads to lower background cleavage on the gels of photo-irradiation experiments (Odom et al., 2000). The oligonucleotides then were dissolved in 50 μl of denaturing gel loading dye (99% formamide (v/v), 1 mM Tris-Cl pH 7.5, and 1 mM EDTA pH 8.0) and heated at 100 °C for 5 minutes to denature any preformed nucleic acid structures. The oligonucleotides were size-purified in 10% (w/v) denaturing polyacrylamide gels and visualized by UV shadowing. The oligonucleotides were excised from the gel and eluted via crush-soak using 25 ml of TE buffer (10 mM Tris-Cl pH 7.5 and 0.1 mM EDTA pH 8.0) at 4 °C. The eluted samples were filtered through 0.2 μm micro filters and then concentrated using 2-butanol extraction until a
volume of 300 μl is reached. The samples were ethanol precipitated (see appendix for exact procedure). The samples were resuspended in 100 μl TE buffer. The concentrations of the oligonucleotides were determined by UV absorption at 260 nm using a Varian Cary 300 UV/Vis spectrophotometer. All oligonucleotides were stored at -20 °C.

2.2.3 C6-Amino DNA Synthesis and Purification

The DNA oligonucleotides were synthesized at the University of Calgary Core DNA Services. The oligonucleotide to be derivatized with AQ was synthesized with a 5’-C6-amino functionality. Upon receipt, the DNA was treated by the following procedure to remove any possible nitrogenous contaminants from the DNA synthesis procedures. The oligonucleotide was dissolved in 200 μl TE and extracted three times with 400 μl chloroform. The aqueous fraction was ethanol precipitated using NaCl and anhydrous ethanol (see appendix for exact procedure). The sample was vortexed, placed in dry ice to freeze and then centrifuged at 4 °C at 16.1 relative centrifugal force(RCF) for 40 minutes. The supernatant was removed and the DNA pellet was washed with 150 μl cold 70% (v/v) ethanol. The DNA pellet was resuspended in 100 μl ddH₂O. The concentrations of the oligonucleotides were determined by UV absorption at 260 nm using a Varian Cary 300 UV/Vis spectrophotometer. The oligonucleotide was stored at -20 °C.

2.2.4 Anthraquinone Coupling of C6-Amino DNA

A 20 μg/μl AQ-NHS ester stock solution was prepared by dissolving AQ-NHS ester in anhydrous DMF. The standard anthraquinone (AQ) coupling reaction was composed of 12 nmol of C6-amino DNA, 75 μl of 100 mM Na2B4O7, 7 μl of 20 μg/μl anthraquinone (AQ) stock solution, and ddH₂O to make a final volume of 100 μl. The
sample was covered by aluminium foil to keep the sample in the dark and was shaken vigorously overnight at room temperature. The sample was then briefly spun-down and the DNA solution was transferred to another 0.65 ml microcentrifuge tube. The sample was then ethanol precipitated with NaCl and anhydrous ethanol (see appendix for exact procedure). The DNA pellet was washed with 150 μl cold 70% (v/v) ethanol. The pellet was fully resuspended in 50 μl of 100 mM triethylammonium acetate (TEAA) (pH 6.85) and 100 μl chloroform. The uncoupled AQ also precipitated and formed a large pellet along with the AQ-DNA. In order to separate the AQ-DNA from the uncoupled AQ, the uncoupled AQ need to be fully dissolved; this is why a two-phase solvent system was used to efficiently separate the AQ-DNA from the uncoupled AQ. The chloroform layer dissolves the uncoupled AQ while the aqueous layer had AQ-DNA. The aqueous layer was extracted two times with 100 μl chloroform and the aqueous fraction was lyophilized under vacuum. The dried DNA was resuspended in 25 μl of 50 mM TEAA (pH 6.85).

![Anthraquinone coupled to DNA through a peptide bond between anthraquinone and the 5’ amino group on C6 linker of DNA.](image)

Figure 2.1 Anthraquinone coupled to DNA through a peptide bond between anthraquinone and the 5’ amino group on C6 linker of DNA.

### 2.2.5 HPLC Purification of Anthraquinone-coupled DNA

The AQ-coupled DNA was purified by high-pressure liquid chromatography (Agilent 1100 series HPLC) using a reverse-phase C-18 column. The solvent flow rate
was continuously at 1 ml/minute and the solvent was heated to 40 °C. The initial conditions were 100% solvent A (20:1 50 mM TEAA (pH 6.85): acetonitrile), changing to 40% solvent B (100% acetonitrile) over 42 minutes with a linear gradient. Following this period, the solvent was changed to 100% solvent B for 10 minutes to flush out and clean the column before reconditioning the column to the original conditions. Typically, uncoupled DNA has a retention time of approximately 7 minutes, while the AQ-DNA has a retention time of approximately 13 minutes. Retention times vary for different sequences depending on the length of the DNA. To ensure that the collected fractions were indeed AQ-DNA a ratio of the area of the absorbance peak at 260 nm was divided by the area of the absorbance peak at 335 nm. AQ-DNA should have a ratio range (60-100) depending on the length and sequence of the DNA (Fahlman and Sen, 2002). The fractions containing the AQ-DNA were lyophilized under vacuum and resuspended in 50 μl ddH₂O. The AQ-DNA concentration was determined by UV absorption at 260 nm using a Varian Cary UV/Vis spectrophotometer. A typical yield of AQ-DNA conjugates ranges from 50-80% of the starting material, depending on the sequence and synthetic batch.

The exact retention times of the DNA oligonucleotide varies depending on its length, but in all cases, the uncoupled DNA elutes from the reverse-phase HPLC column before the coupled DNA. Absorbance spectra were taken at 260 nm (A) and 335 nm (B). DNA absorbs only at 260 nm and AQ absorbs at both 260 and 335 nm, therefore the peak at approximately 7 minutes is the uncoupled DNA and the peak at approximately 13 minutes is the DNA coupled to AQ. The final broad peak at 20 minutes is the uncoupled AQ.
2.2.6 RNA synthesis

All RNA strands were synthesized using T7 RNA polymerase, with a double-stranded template incorporating the T7 RNA polymerase promoter. The template was amplified using the polymerase chain reaction (PCR) with Vent DNA polymerase (New England Biolabs), using standard protocols. Transcription reactions were carried out using purified T7 RNA polymerase (30 µg) and the DNA template (2-5 µM) in 100 µl of transcription buffer (40 mM Tris-HCl (pH 7.9), 26 mM MgCl2, 2.5 mM spermidine, 10 mM DTT, 0.01% Triton X-100, 8 mM GTP, 4 mM ATP and CTP, 2 mM UTP), the reaction proceeding for 2-16 h at 37 °C. Upon completion, the transcription mixtures were treated with deoxyribonuclease I (Fermentas), for 30 min at 37 °C, and the newly synthesized RNAs ethanol precipitated and size-purified by denaturing polyacrylamide gel electrophoresis. Dephosphorylation (for RNA) using calf intestinal alkaline phosphatase (Roche Diagnostics) and 32P-end labeling (for DNA and RNA) using polynucleotide kinase (Invitrogen) were carried out using standard protocols.

2.2.7 5’ End-Labelling

DNA oligonucleotides were 5’ end-labelled using a standard phosphorylating procedure (see appendix for procedure). The DNA was resuspended in 20 µl denaturing gel loading buffer and purified on 10% (w/v) denaturing polyacrylamide gel. The purified DNA was resuspended in 30 µl TE (10 mM Tris-Cl pH 7.5 and 0.1 mM EDTA pH 8.0).

2.2.8 Assembly of DNA or RNA Constructs

DNA or RNA assemblies were formed by annealing stochiometric mixtures of constituent oligonucleotides (1 mM each), including the AQ-DNA and trace amounts of
5’-end labeled oligonucleotide, in 50 mM Tris-HCl (pH 7.5) and 0.1 mM EDTA. The solutions were heated at 90 °C for 30 s, cooled slowly to room temperature (60-90 min), then diluted 4-fold in irradiation buffer (final composition: 50 mM Tris-HCl (pH 7.5), 50 mM NaCl, 1 mM MgCl2, and 0.1 mM EDTA). The samples were incubated for approximately 15 min at room temperature prior to photo-irradiation to initiate charge flow.

2.2.9 Photo-irradiation of the DNA/RNA complexes

DNA/RNA solutions were placed in the wells of an ELISA plate pre-coated with glycogen (1 mg/ml) and placed under a Black-Ray UVL-56 lamp (365 nm) for 60 min at a distance of 2 cm from the bulb. Temperature was maintained by having the ELISA plate in contact with a water bath. Following photo-irradiation, the samples were ethanol precipitated with the assistance of glycogen as carrier. Glycogen is heavy and helps the micromolar concentrations of DNA precipitated to settle down.

2.2.10 Iridium oxidation

The DNA/RNA pellets were dissolved in 50 mM Tris-HCl (pH 7.5) containing 1 mM Na2IrCl6 (Strem Chemicals), and incubated for 1 hour in the dark. The oxidation was then quenched with 1 µl of 0.1 M EDTA, and the treated RNA/ DNA recovered by ethanol precipitation. The samples were then treated with piperidine or aniline, as described below.

2.2.11 G and A > G ladders

The dimethylsulfate (DMS) and diethylpyrocarbonate (DEPC) Maxam-Gilbert sequencing reactions were carried out following standard protocols, followed by ethanol
precipitation to recover the DNA/RNA (see appendix for detailed process). The modified RNA and DNA samples were then treated with piperidine or aniline, as described below.

### 2.2.12 Aniline and piperidine treatment

For the aniline reaction, the ethanol-precipitated DNA/RNA pellet was dissolved directly into an aniline acetate solution (20 µl aniline, 180 µl H2O, and 12 µl of glacial acetic acid) and strand scission preformed at 60 °C for 15 min in the dark. For the W1 workup, each sample was diluted with 80 µl of doubly deionized water (ddH2O) and vacuum dried for 1 h in a lyophilizer. For the W2 workup, the vacuum-dried pellet was dissolved in 200 µl of ddH2O, and was left to vacuum-dry overnight in the lyophilizer. For the piperidine cleavages, the DNA pellets were dissolved in 90 µl of ddH2O and 10 µl of fresh piperidine were added. The solution was incubated at 90 °C for 30 min in the dark and then vacuum dried for 1 h (for the W1 protocol). For the W2 workup, as above, the pellet from the W1 protocol was dissolved in 200 µl of ddH2O, and was left to vacuum dry overnight in the lyophilizer.

### 2.2.13 Denaturing Gel Electrophoresis

The lyophilized DNA was dissolved in denaturing gel loading dye (99% formamide (v/v), 10 mM Tris-Cl pH 7.5, and 0.1 mM EDTA pH 8.0), heat denatured at 100 °C for 5 minutes, cooled to 21 °C, and loaded on 20% (w/v) denaturing polyacrylamide sequencing gels. The gels were analyzed by phosphorimagery using a Molecular Dynamics Typhoon 9410 Variable Mode Imager.
2.2.14 Data Analyses

The densitometry analyses of the sequencing gels were carried out using Molecular Dynamics Image Quant 5.2 software for Windows 2000. The density of the band of interest was assessed as a percentage of the total signal contained within a particular lane. This procedure compensated for any discrepancies of the total radioactive counts in different lanes. The normalized signal for a particular nucleotide obtained in the "dark" reaction negative control lane was subtracted from the signal of the same nucleotide in an irradiated sample lane.
2.3 Results and Discussion

2.3.1 Design of DNA/RNA duplex for charge transfer experiments

To date, hole transfer through RNA/DNA heteroduplexes has been studied. But to our knowledge, charge flow from a tethered photooxidant through purely RNA helices had never been reported. We wished to undertake a systematic comparison of the hole transfer properties (as well as the efficiencies of the water reaction leading to CFQ) at guanines located within a B-type DNA double helix relative to an A-type RNA double helix of identical sequence.

The nucleic acid stems used in our study are shown diagrammatically in Figure 2.2.A and B. In all the experiments, three distinct oligonucleotides were assembled together to generate the duplex for charge flow. In the case of the all-DNA substrate, a 28 nucleotide oligonucleotide was hybridized with an 18 nucleotide and a 10 nucleotide oligonucleotide, where the last oligonucleotide is covalently derivatized at its 5’ terminus to an anthraquinone (AQ) moiety (Fig. 2.2.C). Once hybridisation of all the three oligonucleotides was accomplished, a blunt-ended DNA duplex (incorporating a single nick between the 10 and 18 nucleotide oligonucleotides) was ready for photo-irradiation experiments. The utility of this particular design is that, when required, the 18 and 28 nucleotide DNA oligonucleotides may be substituted with RNA oligonucleotides of the same sequence (while the AQ-derivatized oligonucleotide remains DNA in all cases).

The original idea of the project was conceived by a post-doctoral fellow, Dr. Lucien Junior Bergeron, from the Sen laboratory. He had designed the DNA sequences for constructing the model DNA duplex and had succeeded in getting it to conduct charges. Moreover, preliminary promising results were also achieved with the RNA
duplex that was transcribed from the original DNA template that constituted the model duplex. When I joined the laboratory as a graduate student, I started working with him on this project. Since the results of experiments on RNA were promising, Dr. Bergeron was expected to continue further with the characterisation of the duplexes by introduction of variations like replacing complementary base pairs with mismatched base pairs, addition of 5’ and 3’ single strand overhangs etc. As he had to leave, I took over the responsibility of continuing with those experiments till reproducible results were achieved, that resulted in a publication with the title “A guanine-linked end-effect is a sensitive reporter of charge flow through DNA and RNA double helices” in 2008 in the Biochimie journal. All the experiments and their corresponding gel figures in the paper (and in this chapter of my thesis) were performed by me (with the exception of Figure 4 of the publication which is the same as Figure 2.5 of chapter two of my thesis).

The nucleic acid duplex was designed taking three important features into consideration: (a) the DNA sequence does not incorporate more than one consecutive A-T base pair. As A-T base pairs are known to inhibit charge flow, the GC base pairs acting as the primary charge carriers should not be distanced from each other more than one A-T base pair; (b) the requirement of T7 RNA polymerase system for the oligonucleotide to have a particular sequence of 5’GGA... 3’ at the 5’ end of all the in vitro transcribed RNA oligonucleotides (Milligan and Uhlenbeck, 1989). The DNA oligonucleotides coding for the transcribed RNA with this specific sequence at their 5’ ends, gives the highest yield of the transcription product, hence the need for such a sequence (Milligan and Uhlenbeck, 1989). (c) a conventional GGG hole trap need to be placed (shown with a bracket in Fig. 2.2.B) three base pairs away from the end of the duplex that is distal to the
AQ moiety. This particular last feature of the duplex resulted in a 5’-GGAGGG.... sequence at the distal end of the duplex making it unique and different from most duplexes used by other research groups for hole conduction studies. Stretches of guanines at the extreme distal end has not been used in most of the duplexes experimented with, so far with charge transfer (reviewed in Núñez et al., 1999 and in Giese, 2002).

Within a double helix, the level of damage caused by charge flow quenching technique (CFQ) at any given guanine is proportional to the time which the mobile radical cation resides at that guanine base. All densitometry signals were calculated as a percent based on the counts of the band of interest divided by the total counts of that lane. The signal for charge transfer dependent oxidative damage of a particular base was obtained by subtracting the background percent damage of the same corresponding base from the dark reaction.
Figure 2.2  (A) Schematic diagram for the construction of the model double helix, WT. AQ indicates the covalently appended anthraquinone. The oligodeoxynucleotide linked to the AQ is shown in gray. The predicted charge transfer detector sequence, GGG, is shown as CT detector. The electron flow direction is indicated by (e to +). (B) Sequences of the oligodeoxynucleotides used to make up the DNA WT duplex. The site for introduction of the different base pair/mismatch combinations is shown (at position 12) with a bracket. The DNA nucleotides linked to the AQ are boxed. The 28-nt oligonucleotide was $^{32}$P-labeled at its 5'-end. (C) Anthraquinone covalently couples to the 5'-end DNA oligonucleotide with a six carbon link.
2.3.2 Comparison of piperidine and aniline cleavages of DNA

Since hole-conduction-related guanine oxidation is the most convenient way of monitoring charge flow through DNA and RNA helices, there is an inherent problem in monitoring relative charge flow efficiencies in DNA and RNA, because RNA cleavage at oxidized guanines requires a relatively mild workup with acidified aniline, whereas DNA is traditionally treated with 10% v/v piperidine base at 90 °C. Since the overall β-elimination chemistries resulting in strand cleavage are similar but not identical for both the cases (reviewed in Burrows and Muller, 1998; Maxam and Gilbert, 1980; Peattie, 1979), we investigated the utility of aniline for effecting DNA strand cleavage, relative to the standard protocol using piperidine. Such a partially oxidized DNA is treated with an iridium (IV) compound to enable a further one electron oxidation (Muller et al., 1998) and the production of guanine oxidation products that are highly labile to hot aqueous piperidine. Such a treatment has been found to generate modestly higher levels of charge-flow generated DNA cleavage.

While aniline treatment can indeed substitute for piperidine treatment for DNA cleavage, it shows a level of complexity of the cleavage products generated. Figure 2.3 shows that two slightly different protocols: (a) where, following aqueous aniline/piperidine workup, the cooled DNA solutions were quickly lyophilized to dryness and analyzed (named protocol “W1”), and (b) where the aniline/piperidine treated DNA were lyophilized twice, over an extended period, where the DNA pellet was solubilised with a fresh aliquot of deionised water for a second round of lyophilisation (named protocol “W2”).

Figure 2.3.A and B shows that piperidine treated DNA produced guanine cleavage bands that could be assigned unambiguously, when subjected to either W1 or
W2 protocols. The aniline workup with the W1 and W2 protocols gave rise to bands that had slightly different mobilities. In Fig. 2.3.A and B, lanes 1 (worked up with W1) and 3 (worked up with W2) show ‘‘dark’’ samples, with no expected charge flow-related cleavage, whereas lanes 2 and 4 show the irradiated duplex (i.e. subjected to charge flow) worked up, again, with W1 and W2, respectively. Lanes 5 (W1) and 6 (W2) show dimethylsulfate (DMS)-generated Maxam-Gilbert guanine sequence ladders, and lanes 7 (W1) and 8 (W2) show diethylpyrocarbonate (DEPC)-generated A > G ladders. In lanes 2 and 4 of Fig. 2.3.A and B one can see characteristic charge-transfer-generated guanine bands (at the GGG motif, guanines 4-6) relative to the ‘‘dark’’ lanes, 1 and 3.

But the most striking observation from the aniline workup (Fig 2.3.A) is the three remarkably intense bands (labelled 1a, 1b, and 1c), that arise as a consequence of charge transfer from the 5’-most nucleotide(s) of the 5’-32P-labeled 28 nucleotide duplex oligonucleotide. Upon careful identification, it appeared that the intense bands 1a, 1b, and 1c did indeed arise from the 5’-most guanine of the sequence (cumulatively, these bands account for 20% and 16% of the total counts in lanes 2 and 4, Fig. 2.3.A, respectively).

Piperidine workup (Fig. 2.3.B), however, did not generate the same multiplicity of gel bands as the aniline workup (Fig. 2.3.A). With piperidine, the W1 and W2 protocols generated only two products of identical gel mobility, labelled as 1’ and 1”. This could be attributed to the 5’-most G of the 28-nucleotide sequence. The cleavage intensities of bands 1’ and 1”, were not as high as those found for 1a, 1b, and 1c (Fig. 2.3.A). When run side-by-side, the gel mobilities of 1’ and 1” were notably greater than those of 1a, 1b, or 1c, suggesting that the three aniline-generated products were probably aniline adducts.
of DNA breakdown products seen with piperidine treatment. Mass spectroscopy of the bands 1a, 1b, 1c, 1’ and 1” could not determine their chemical identities.

We thus hypothesized that these end-effect products (particularly 1a, 1b, and 1c from the aniline workup) might sensitively “report” successful charge flow through given DNA (and, RNA) double helices. A series of experiments were done to determine (a) the sequence requirements for these breakdown products; (b) their fidelity and reliability as indicators of successful charge flow; and (c) whether they are able to report charge-transfer through DNA as well as RNA.
Figure 2.3  DNA cleavage bands from aniline (A) and piperidine (B) workups of charge-flow-oxidized WT DNA. The data explore the relative effects of Protocol 1 (W1) versus Protocol 2 (W2) for the aniline/piperidine workups (see text). Lanes 1 and 3: dark (no charge flow induced) WT duplex, worked up with the W1 and W2 protocols, respectively. Lanes 2 and 4, light-irradiated (charge flow induced) WT duplex, worked up with the W1 and W2 protocols, respectively. Lanes 5 and 6, dimethylsulfate (DMS)-generated guanine ladders, worked up with the W1 and W2 protocols, respectively. Lanes 7 and 8, diethylpyrocarbonate (DEPC)-generated A > G ladders, worked up using the W1 and W2 protocols, respectively. The lines linking bands in the DMS and DEPC ladders shown in lanes 5-8 (A) indicate tentative correlations of sequence bands generated by the W1 and the W2 workup conditions.
2.3.3 To test that the products 1a, 1b and 1c are produced as a result of charge flow.

We wished to test whether the end-effect products 1a, 1b, and 1c, were actually products of charge conduction through the DNA duplex (Fig. 2.4). For charge conduction to occur in our experimental system, the following steps/reagents are required: (i) the covalently appended AQ photooxidant; (ii) UV light irradiation (356 nm); (iii) workup with hot aniline/piperidine. A fourth step (iia), carried out between steps (ii) and (iii), and is the treatment of the partially oxidized DNA with an iridium (IV) compound (to further oxidize the guanine oxidation products and render them more susceptible to aniline / piperidine treatment) (Muller et al., 1998).

We wished to test if the end-effect bands were produced only when, (i), (ii), and (iii) all the three conditions had been satisfied, and also whether treatment (iia) generated them adventitiously, even in the absence of aniline workup. In Fig. 2.4, lanes 1 and 3 showed the results of DNA duplexes that lacked AQ (i) and light (ii), but were treated with iridium (iia), followed by aniline (iii) under the W1 (lane 1) and W2 (lane 3) conditions, respectively. Lanes 2 and 4 contained the DNA duplexes lacking AQ (i), but treated with light (ii), iridium (iia), and aniline (iii) under W1 (lane 2) and W2 (lane 4) conditions, respectively. In none of the lanes 1-4, the end-effect products are higher intensity compared to the background. Lanes 5 and 7 represent W1 and W2 aniline workups of DNA duplexes subjected to (i), (iia), and (iii), but not irradiated with light (ii); whereas, lanes 6 and 8 show W1 and W2 products of DNA duplexes treated to all four conditions. Lanes 9 and 10 show Maxam-Gilbert guanine ladders generated by dimethylsulfate treatment followed by aniline workup under W1 (lane 9) and W2 (lane 10). Significantly, it is only in lanes 6 and 8 that the end-effect bands are notably present,
consistent with their appearance as a consequence of charge flow through the DNA. Although lanes 5 and 7 indicate that treatment of AQ (i) - derivatized duplexes with iridium (ii) followed by aniline (iii) is not sufficient (in the absence of light) to generate the end-effect bands, we tested for the possibility that the iridium treatment alone, even in the absence of aniline workup, might contribute to the end-effect bands. Lanes 11-14 show the AQ derivatized duplex not treated to light, iridium, or aniline (lane 11); not treated with iridium or aniline but irradiated (lane 12); treated with iridium but not to light or with aniline (lane 13); and, treated with iridium and light, but not with aniline (lane 14). None of the above lanes show end-effect products at higher than background levels. We did additional controls (not shown in Fig. 2.4) in which the iridium treatment was eliminated, but the AQ-containing duplexes were irradiated and worked up with aniline/piperidine in the usual way. These duplexes showed end-effects exactly comparable to those shown by duplexes treated with iridium. For the rest of our investigation, we followed the W1 workup protocol using aniline, exclusively, since its primary end-effect product was the substantial and well-defined DNA fragment, 1a.
Control experiments to investigate whether the end-effect bands, 1a-1c, are indeed charge-transfer generated, requiring AQ, light, as well as treatment with aniline. Lanes 1 and 3 show the results of DNA duplexes that lacked AQ and light, but were treated with iridium, followed by aniline under the W1 (lane 1) and W2 (lane 3) conditions, respectively. Lanes 2 and 4 show DNA duplexes lacking AQ, but treated with light, iridium, and aniline under W1 (lane 2) and W2 (lane 4) conditions, respectively. Lanes 5 and 7 represent W1 and W2 aniline workups of DNA duplexes with anthraquinone, treated with iridium and aniline, but without light/charge-flow, whereas, lanes 6 and 8 show W1 and W2 products of charge-flow DNA duplexes treated to all four conditions. Lanes 11-14 show the AQ-derivatized duplex not treated to light, iridium, or aniline (lane 11); not treated with iridium or aniline but irradiated (lane 12); treated with iridium but not to light or with aniline (lane 13); and, treated with iridium and light, but not with aniline (lane 14). Lanes 9 and 10 are the DMS-generated guanine ladders for W1 and W2 respectively.
2.3.4 Does the observed end-effect require a terminal G-C base pair?

Two hypotheses for the origin of the end-effect were that (a) the six consecutive purines (five of which are guanines) at the end of the helix represent a necessary and deep charge sink for generating the end-effect, and (b) residues at the end of the helix and, in particular, the terminal base pair, enjoy a large exposure to the solvent, making them unusually prone to the oxidative ‘water reaction’ which in turn generates the large end-effect. To test whether a terminal G-C base pair was absolutely required, we constructed a series of variants (D1, D2, and D3) on our model helix (WT) (Fig.2.5.A), which added one, two, and three A-T base pairs, respectively, to the end of helix past the six-purine stretch. The fourth lane (right lane under WT) in Fig 2.5.B shows the large end-effect bands from the unmodified duplex, WT. Following the 1a band from lane 4 to its counterparts in lanes 6, 8, and 10 (generated from the D1, D2, and D3 duplexes, respectively) one finds two key features: (a) there is the expected monotonic decrease in gel mobilities (G1 is responsible for the end-effect in the WT duplex, but G2 in D1, G3 in D2, and G4 in D3), which run as 1, 2, 3, and 4 nucleotide fragments of DNA, respectively; and (b) there is a corresponding reduction in the intensities of these various 1a bands (the 1a band of the WT duplex accounts for 9.6% of the total counts, but 6.0% of D1, 3.4% of D2, and 2.1% of D3, respectively). By comparison, the cleavage data at the intrahelical guanines G4, G5, G6, and G7 (the 5’ most guanine in each GGG stretch) in, respectively, the WT, D1, D2, and D3 duplexes show cleavages at 0.6%, 1.4%, 1.8%, and 1.4%, typical cleavage levels resulting from charge flow (reviewed in Núñez et al., 1999; Giese, 2002; Schuster, 2000). Indeed, the large 1a band seen with the WT duplex (9.6% of the total counts in the lane) reduces, in the D3 duplex, to a low-level (2.1%) DNA cleavage for the corresponding, now intra-helical, guanine. Given that the purine-
rich charge sink (GGAGGG) is invariant within the above four duplexes, the above results suggest that hypothesis (b), namely, the high efficiency of water reaction at or close to a helix end, is mainly responsible for generating the end-effect.
Figure 2.5  (A) The end sequences of duplexes used for the data shown in (B). (B) Denaturing 20% polyacrylamide gel of the cleavage products from duplexes WT, D1, D2, and D3 respectively. Lanes 1, 2, 3 and 4 correspond to guanine ladders for duplexes D1, WT, D2, and D3 respectively. For each pair of lanes labelled WT, D1, D2, and D3, the left lane represents non-irradiated DNA (which have not been subjected to charge flow), whereas, the right lane represents DNA samples that have been irradiated and thus subjected to charge flow.
2.3.5 The effect of 5’ and 3’-end single-strand overhangs

The above experiments indicated that the end-effect could be observed provided there was a sizable charge sink at or close to the end of the double helix, with a G-C base pair no further than three base pairs away from the helix terminus. We now investigated whether the end-effect required the double helix to have a blunt end. The following variants (shown in Fig. 2.6.A) of the WT duplex were constructed: D4 to D7 maintained the sequence of the WT duplex but extended the C-rich strand by 1-4 nucleotides; duplex D8 had a blunt end, with an extra GC base pair added to the WT sequence, to make one of the two strands 5’GGGAGGG…3’; D9 and D10 had the C-rich strand of D8 shortened by 1 and 2 nucleotides, respectively, leading to the end-effect generating terminal guanine being localized in a single-stranded overhang. Fig. 2.6.B shows that the duplexes D4-D8 generate end-effects comparably to the WT duplex; by contrast, the end-effect is severely diminished in D9 and even more in D10. This same trend was observed when workup was carried out with piperidine rather than aniline (data not shown). These results appear, at first sight, to be in contradiction to those of Kan and Schuster (Kan and Schuster, 1999) who studied the propagation of the radical cation through extended single-stranded tails attached to duplexes. Those authors reported that charge-flow generated guanine oxidation could clearly be observed (via strand-cleavage following hot piperidine treatment) in single-stranded GG doublets that were 1-3 nucleotides away from the duplex end (with the striking feature that the two guanines in such a doublet were oxidized to an equivalent extent, unlike doublets within a duplex, consistent with the single-strand oxidation product being diaminooxazolone rather than the 8-oxoguanine formed in duplexes). In our D9 and D10 constructs, however, the end-effect was found to be severely diminished relative to the parent duplex, whether workup was carried out
with aniline or with piperidine. The source of the discrepancy between our results with those of Kan and Schuster (Kan and Schuster, 1999), though not immediately obvious, may lie in the differences in our constructs. D9 and D10 terminate in very short overhangs (1 and 2 nucleotides, respectively) that terminate in blocks of guanines, whereas the constructs of Kan and Schuster typically contained larger overhangs that did not terminate in blocks of guanines. Undoubtedly, eventual assignment of the chemical identities of our end-effect products will throw light on these discrepancies.
Figure 2.6  (A) The end sequences of duplexes used for the data shown in (B). (B) Denaturing 15% polyacrylamide gel of the cleavage products from duplexes WT and D4-D10. For each pair of lanes labelled as WT, D4, etc., the left lane represents non-irradiated DNA (which has not been subjected to charge flow), whereas the right lane represents DNA that has been irradiated and thus subjected to charge flow. Lanes DMS represent guanine ladders.
2.3.6 Is the end-effect sensitive to base pair mismatches within the double helix?

Charge flow through a DNA double helix has been shown to be sensitive to the presence of mismatches and non-canonical base pairs within the duplex. A comprehensive work on this subject, by Bhattacharya and Barton (Bhattacharya and Barton, 2001) studied charge flow induced by a ruthenium photo-oxidant in a short double helix. The authors compared the ratio of oxidative damage at distal guanines versus proximal guanines on opposite sides of a designated base pair, which was varied systematically to give all 16 canonical base pairs and mismatches (this mutable base pair, X-Y, was placed within the sequence 5’...GACAXTGTG...3’, with ‘‘X’’ base pairing to ‘‘Y’’ on the complementary strand, and both X and Y being varied to be A, G, T, or C). The authors concluded the ability of canonical base pairs (at the designated position) to support charge flow in the following order: CG > GC » AT, TA; and, mismatches following the order: (GC) ≈ GG, GT, TG, AG, GA > AA > CC, TT, CA, AC, CT, TC.

To test whether the end-effect in our aniline-treated DNA system gave similar or discrepant results to the above, we systematically incorporated all 16 base pair/mismatch combinations at position 12 of the WT duplex (Fig. 2.2.B). Our placement of the base pair/mismatch combinations was in a sequence context (5’..GTCTXACGT..3’) somewhat different from that used by Bhattacharya and Barton (Bhattacharya and Barton, 2001), but similar in as much as the mismatch site was flanked by A-T base pairs on either side (with a G-C base pair beyond each A-T base pair). Fig. 2.7 shows the appearance (or non-appearance) of the end-effect bands for each of the 16 possible base pair/mismatch combinations in our mutated duplexes. Even viewed qualitatively, the following order of end-effect (gel bands 1a, 1b, 1c) production is manifest: CG ≈ GC » AT, TA; and, (CG) ≈ GG, GT, TG, GA > AA > CC, TT, CA, AC, CT, TC, AT and
TA. In this list of all sixteen pairs of bases, the base on the left was in the AQ strand, while the base on the right was in the radio-labelled strand. These end-effect data are in excellent agreement with the order of the distal/proximal guanine oxidation ratios reported by Bhattacharya and Barton (Bhattacharya and Barton, 2001), even though the two studies were carried out on duplexes with different sequences. In further agreement with those authors (Bhattacharya and Barton, 2001), we found that canonical A-T and T-A base pairs placed within our sequence context, giving rise to three A-T/ T-A base pairs in a row, severely reduced charge flow to the distal end of the helix (it has been shown elsewhere (Giese et al., 2001) that three A-T/T-A base pairs are not inherently insulating; for instance, three adenines in a row permit charge-flow via the phenomenon of “A-hopping”). Overall, comparing the distal/ proximal damage ratio approach to ours, it appears that the end-effect data provide a direct and unambiguous visual readout that is, if anything, more sensitive, yielding numbers ranging between 13.8% (for C-G) and 0.77% (for the T-T mismatch), whereas the distal/ proximal ratios ranged between 2.05 (for C-G) and 0.15 (for T-C) (Bhattacharya and Barton, 2001).

Two conclusions can be reached from our results: (a) some of the sequence mismatches introduce conformational/ dynamical changes to the double helix to clearly affect/ inhibit charge flow; and (b) the results provide unambiguous proof that the charge flow effect seen in a given duplex is not the consequence of cross-interaction of that duplex with an AQ moiety tethered to a different duplex. An interesting feature of the charge-flow patterns shown in Fig. 2.7 is that in those duplex constructs where charge-flow is blocked owing to the presence of a mismatch (such as AA or CA), enhanced
damage is observed at the AQ-proximal G22 and G19 residues, relative to damage at those residues in the CG-containing parent duplex.
Figure 2.7 20% denaturing polyacrylamide gels showing the influence on the end-effect bands (1a-1c) of substituting the full set of base pair and mismatch permutations at position 12 of the WT duplex (refer Fig. 2.2.B). The lane at the far left (DMS) shows a guanine ladder.
2.3.7 Does an RNA double helix generate the same end-effect?

A key goal of this study was to determine whether an RNA/ RNA duplex conducted charge comparably to a DNA/DNA or DNA/RNA duplex. To carry out an experiment on an RNA duplex, we used the simple strategy of utilizing a modified WT duplex (Fig. 2.2.B), in which the AQ-derivatized 10 nucleotide oligonucleotide was DNA, but the 18 and 28 nucleotide oligonucleotides were now RNA. Fig. 2.8 shows the guanine-damage patterns from charge conduction experiments carried out with the DNA/ DNA and RNA/ RNA duplexes (lanes 2 and 4 were un-irradiated controls; lanes 3 and 5 show samples irradiated to set up charge flow). The most notable result was the appearance of the end-effect, which is of comparable intensity for both the DNA (band 1a) and RNA (band 1) duplexes. The mobility difference observed for those bands is only due to higher molecular weight of RNA compare to DNA. However, two striking differences were also noted: (a) regarding the end-effect itself, DNA generated the characteristic three products, 1a, 1b, and 1c; however, RNA generated only one product (band 1), of comparable gel mobility to product 1a from DNA: (b) The RNA/ RNA and DNA/ DNA double helices showed different intra-helical guanine oxidation patterns. Whereas, DNA showed the characteristic guanine-damage patterns associated with charge transfer, with notable damage at the GGG site (mainly at positions 4 and 5), in RNA there appeared to be relatively little intra-strand cleavage at those positions in the irradiated lane (lane 5) relative to the dark lane (lane 4). Possibly, this could reflect the differences of water access (poorer access in RNA than in DNA) via the structurally different major and minor grooves of A-type (RNA) relative to B-type (DNA) helices, or the formation of different oxidation products from these intra-helical guanines in DNA and in RNA. Interestingly, a prior study (Fernandez-Saiz et al., 1999) on short RNA
hairpins irradiated in the presence of free, dissolved AQ reported the absence of charge-flow related RNA cleavage at intra-helical guanines (as opposed to those in the loop). Likely, these prior observations, as well as ours, indicate one of two possibilities: (a) the actual lack of significant charge-flow generated oxidation of stem (intra-helical) guanines in RNA, relative to DNA; or (b) the formation of distinct oxidation products in RNA versus DNA, with concomitantly different susceptibilities to aniline cleavage. Our preliminary investigations of these possibilities suggest that the introduction of mismatches within the RNA double helix do modulate or eliminate the end-effect.
Figure 2.8 Comparison of patterns of charge flow through DNA and RNA WT duplexes (with thymine in DNA sequences replaced by uracil in the RNA sequences). The short AQ-derivatized DNA strand is conserved in both systems. Lanes 2 (DNA) and 4 (RNA) represent non-irradiated DNA (which have not been subjected to charge flow), whereas lanes 3 (DNA) and 5 (RNA) represent samples that have been irradiated and thus subjected to charge flow. Lanes 1 and 6 represent G ladders for the DNA and the RNA, respectively.
2.4 Conclusion

To the best of our knowledge, this was the first demonstration of charge flow through a RNA-RNA double helix. Our experiments have shown that DNA and RNA double helices with G-rich termini are capable of generating a massive oxidative-end effect and hence serve as a sensitive reporter of charge flow. This end effect is highly sensitive to specific base-pair sequences and mismatches.

The reasons for the origin of end-effect could be explained by various prior studies:

(i) NMR experiments have demonstrated the reality of end-fraying in DNA double helices (Patel et al., 1975).

(ii) In the absence of base pairing, the G\(^{+}\) radical cation can undergo irreversible deprotonation to give the guanine radical (G\(^{+}\)), which is being preferentially oxidized to 2, 2-diaminooxazolone instead of 8-oxoguanine (Kan and Schuster, 1999; Cadet et al., 1994; Kino et al., 1998).

(iii) The DNA sites containing 2, 2-diaminooxazolone lesions are known to cleave with noticeably higher efficiency upon piperidine treatment, when compared to the sites containing 8-oxoguanine (reviewed in Burrows and Muller, 1998; Muller et al., 1998).
2.5 Future directions

The magnitude of the end-effect and its sensitivity to single mismatches within the DNA double helices could indicate that it could be exploited as a reporter for monitoring gene mutations.

The charge flow properties of complex nucleic acid assembly such as three-way (Santhosh and Schuster, 2003) and four-way junctions (Odom and Barton, 2001; Fahlman et al., 2002), multi-helix DNA constructs incorporating ligand-binding aptamers have already been studied (Fahlman and Sen, 2002; Sankar and Sen, 2004). The Sen laboratory studied the charge flow and quenching patterns of a complexly folded nucleic acid assembly (the 8-17 deoxyribozyme, bound to its DNA pseudosubstrate), which provided some interesting structural information about it, including the identification of helical stacking preferences, segmental folding properties, and exposure of segments as well as individual nucleotide residues to surrounding solvents (Leung and Sen, 2007). As an analogy, we could also assume that study of charge flow properties through folded RNAs, ribonucleoproteins, and RNA-small molecule complexes could likewise provide important structural information about them. The end-effect “reporter” could also be thought to significantly help in charge flow studies, of not only DNA and RNA folded structures, but their conformational transitions as well.

This tool could also be applied to study nucleoprotein interactions, such as the binding of transcription factors and DNA modifying enzymes. Interesting structural information, such as the changes of stacking interactions and the degree of solvent exposure of the bases, could be obtained about the protein binding site of the DNA with and without the bound protein. Studies carried out on RNA, RNA-DNA heteroduplexes
and chimeras have shown that these molecules can conduct charge (Odom and Barton 2001; O’Neill and Barton 2002). This will also enable further development of this method to study ribozymes and ribonucleoprotein structures.
Chapter 3:
Structural Studies of DNA duplexes containing Trinucleotide Repeats by Photo-induced Charge Flow

3.1 Introduction

Trinucleotide repeats (TNR) are members of a common class of DNA termed as microsatellites, which are sections of chromosome composed of a 1–6 base pair DNA motif that is repeated many times in a head to tail configuration. For instance, a common microsatellite is the dinucleotide repeat motif (CA)$_n$, where $n$ indicates the number of repeat copies. Thus, a genetic locus of (CA)$_5$ would be composed of the sequence CACACACACACACA. Microsatellites are widely dispersed throughout the human genome and other diverse organisms. An interestingly feature of most microsatellites is polymorphism in their repeat copy numbers at any particular genetic locus. Such differences in length are a result of rare mutations that reveal themselves as small instabilities, increases or decreases of 1 or 2 repeat copy numbers over the evolutionary time course. Severe familial diseases result from expansion of specific TNR loci well beyond their normal polymorphic range (Kovtun et al., 2001) (Table 3.1).
Several short tandem DNA repeats are polymorphic and hence their repeat copy numbers are unstable in all genomes. At several loci within the human genome, this repeat instability is associated with diseases (Cleary and Pearson, 2003; Lenzmeier and Freudenereich, 2003). Expansion of trinucleotide (triplet) repeats is the molecular cause of at least 18 human neurological diseases (repeat diseases), including myotonic dystrophy 1 (DM1), Huntington’s disease (HD), and a number of spinocerebellar ataxias (SCAs) (Orr and Zoghbi, 2007). Repeat diseases are characterized by the expansion of a disease-specific triplet repeat tract beyond a threshold of about 25–35 repeats to a length that has pathologic consequences, often involving neuronal death in disease-specific regions of the brain. The discovery of this unique category of human diseases raised three basic questions to the researchers, such as:

1. What is the molecular mechanism underlying triplet repeat instability?

---

Table 3.1 Genomic loci affected by trinucleotide repeat instability.

<table>
<thead>
<tr>
<th>Locus</th>
<th>Repeat</th>
<th>Normal</th>
<th>Disease</th>
<th>Phenotype</th>
<th>Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>Huntington disease</td>
<td>CAG</td>
<td>11–34</td>
<td>35–121</td>
<td>Neurodegeneration</td>
<td>Coding</td>
</tr>
<tr>
<td>Kennedy’s disease</td>
<td>CAG</td>
<td>11–33</td>
<td>38–66</td>
<td>Neurodegeneration</td>
<td>Coding</td>
</tr>
<tr>
<td>DRPLA</td>
<td>CAG</td>
<td>8–35</td>
<td>40–100</td>
<td>Neurodegeneration</td>
<td>Coding</td>
</tr>
<tr>
<td>Spinocerebellar ataxia 1</td>
<td>CAG</td>
<td>6–40</td>
<td>40–83</td>
<td>Neurodegeneration</td>
<td>Coding</td>
</tr>
<tr>
<td>Spinocerebellar ataxia 2</td>
<td>CAG</td>
<td>15–29</td>
<td>35–59</td>
<td>Neurodegeneration</td>
<td>Coding</td>
</tr>
<tr>
<td>Spinocerebellar ataxia 3</td>
<td>CAG</td>
<td>14–37</td>
<td>68–84</td>
<td>Neurodegeneration</td>
<td>Coding</td>
</tr>
<tr>
<td>Spinocerebellar ataxia 6</td>
<td>CAG</td>
<td>4–16</td>
<td>21–27</td>
<td>Neurodegeneration</td>
<td>Coding</td>
</tr>
<tr>
<td>Spinocerebellar ataxia 7</td>
<td>CAG</td>
<td>4–35</td>
<td>37–200</td>
<td>Neurodegeneration</td>
<td>Coding</td>
</tr>
<tr>
<td>Spinocerebellar ataxia 8</td>
<td>CTG</td>
<td>13–89</td>
<td>107–127</td>
<td>Neurodegeneration</td>
<td>Noncoding</td>
</tr>
<tr>
<td>Fragile X (A)</td>
<td>CGG</td>
<td>6–54</td>
<td>&gt;200</td>
<td>Mental Retardation</td>
<td>Noncoding</td>
</tr>
<tr>
<td>Fragile X (E)</td>
<td>CGG</td>
<td>6–25</td>
<td>&gt;200</td>
<td>Mental Retardation</td>
<td>Noncoding</td>
</tr>
<tr>
<td>Fragile X (F)</td>
<td>CGG</td>
<td>6–29</td>
<td>&gt;200</td>
<td>Fragile site</td>
<td>Noncoding</td>
</tr>
<tr>
<td>Fragile site 16A</td>
<td>CGG</td>
<td>16–49</td>
<td>&gt;200</td>
<td>Fragile site</td>
<td>Noncoding</td>
</tr>
<tr>
<td>Jacobsen’s syndrome</td>
<td>CGG</td>
<td>8–14</td>
<td>&gt;100</td>
<td>Fragile site</td>
<td>Noncoding</td>
</tr>
<tr>
<td>Friedreich’s ataxia</td>
<td>GAA</td>
<td>7–22</td>
<td>&gt;200</td>
<td>Neurodegeneration</td>
<td>Noncoding</td>
</tr>
<tr>
<td>Myotonic dystrophy</td>
<td>CTG</td>
<td>5–27</td>
<td>&gt;50</td>
<td>Neurodegeneration</td>
<td>Noncoding</td>
</tr>
<tr>
<td>SEF2-1</td>
<td>CAG</td>
<td>10–2100</td>
<td></td>
<td>Normal</td>
<td>Noncoding</td>
</tr>
<tr>
<td>ERDA1</td>
<td>CAG</td>
<td>7–93</td>
<td></td>
<td>Normal</td>
<td>?</td>
</tr>
</tbody>
</table>

2. How does expansion of such repeat sequences cause human diseases?

3. And most importantly, could these diseases be treated or prevented (Lin et al., 2009)?

In the last 15 years, several research groups have tried to address these questions. Studies done using model systems of bacteria, yeast, flies, mammalian cells and mice have provided important insights into such molecular processes that destabilize the repeats, yet the actual mechanism is still unclear (Lin et al., 2009). The expansion in them has been reported to arise by replication slippage on either the leading (Iyer and Wells, 1999) or the lagging strand (Kang et al. 1995; Freudenreich et al., 1997; Miret et al., 1998; Schweitzer and Livingston, 1998), homologous recombination (Cemal et al., 1999), gene conversion (Freudenreich et al., 1998; Richard et al., 2000), double-strand break repair (Cemal et al., 1999; Freudenreich et al., 1998; Richard et al., 2000), and base excision repair (Lyons-Darden and Topal, 1999). Hence it could be stated that DNA replication, DNA repair, recombination and transcription would have significant contribution in destabilising these repeats. The unstable repeats associated with diseases have been found in both germline and somatic tissues, but the complexity of tissue specific patterns makes any straightforward correlation with either of replication, transcription or repair, impossible (Cleary and Pearson, 2003; Lin et al., 2006).

DNA triplets repeats are present in many genes that have been associated with neurodegenerative disorders (Smith et al., 1995). Seven out of all the sixty-four possible triplet repeats have been associated with human disorders. Four of these sequences – CTG, CAG and CGG/ CCG have proved to be involved in pathogenesis of Myotonic Dystrophy, Huntington’s disease and Fragile-X syndrome respectively (Mandel, 1994;
Pearson and Sinden, 1996; Reddy and Housman, 1997). The discovery of the association between trinucleotide repeat expansions and several hereditary neurological disorders opens doors for disease diagnosis and treatment (Caskey et al., 1992; Willems PJ. 1994). Very often, these neurodegenerative diseases exhibit an increase in repeat unit length, a phenomenon termed dynamic mutation, and also a related increase in severity and/or decrease in the age of onset of the disease with successive generations (Smith et al., 1995). The inheritance pattern of such disorders are typically characterised by a progressive deterioration over generations, as the repeat numbers keep increasing (Lin et al., 2009). There has been intensive research focused on determining the functional role played by such nucleic acid repeat sequences in the mutational mechanisms by which certain inherited human disorders get transmitted. For instance, the DNA trinucleotide repeat 5'- (CAG) n has been found in the open reading frame of gene sequences that are associated with the onset of at least six neurodegenerative disorders (Orr et al., 1993; LaSpada et al., 1991; Koide et al., 1994; Burke et al., 1994; Kawaguchi et al., 1994). Myotonic dystrophy shows a correlation to increased copies of CTG repeats that are located in a 3’ untranslated region of the MD gene (Mahadevan et al., 1992). Another class of triplet expansion diseases are characterized by CCG/CGG triplet repeats, which are associated with hypermethylation of CG islands adjacent to the repeating sequence (Oberle et al., 1991; Jones et al., 1995). This induces formation of fragile sites, such as FRAXA and FRA16A that are found respectively in the genes of X syndrome (Oberle et al., 1991) and Jacobsen syndrome (Jones et al., 1995).

The mutational mechanism behind the expansion of trinucleotide repeats is poorly understood. A significant question being asked is whether these triplet repeats would
associate in an unconventional structure capable of serving as a basic structural motif in formation of stable or transient in vivo molecular assemblies (Smith et al., 1995). These observations demonstrate the need for understanding the structure-function relationships of such DNA repeats.

3.1.1 Structural information obtained from prior biophysical studies on trinucleotide repeat sequences.

Fragile X triplets (GGC)\(_n\)·(GCC)\(_n\) are located upstream of the FMR-1 gene (Verkerk et al., 1991). It has been more than a decade that the three-dimensional structures of the fragile X triplet repeats, (GCC)\(_n\) and (GGC)\(_n\) have been elucidated by NMR spectroscopy (Mariappan et al., 1996). In various solution conditions that were close to physiological pH, (GCC)\(_5\–7\) strands were seen to form slipped hairpins with a 3’ overhanging cytosine. The slipped hairpins of (GCC)\(_n\) strands exhibit the following structural features: (i) maximization of Watson–Crick G·C pairs; (ii) formation of C·C mispairs at the CpG steps in the stem; (iii) C2’-endo, anti conformations for all the nucleotides (Mariappan et al., 1996). Electrophoretic mobility of native gels show that (GGC)\(_n\) strands form both hairpin and mismatched duplex structures in 10 – 150 mM NaCl (pH 6–7) for \(n < 10\), but when \(n > 10\), they are only shown to form hairpin structures (Mariappan et al., 1996).

On the other hand, NMR analysis of [(GGC)\(_n\)]\(_2\) strands, for \(n = 4–11\) has shown that the most prevalent state was the duplex state. NMR analyses of [(GGC)\(_n\)]\(_2\) duplexes for \(n = 4–6\) show the presence of Watson–Crick G·C and mismatched G\(_{anti}\)·G\(_{syn}\) pairs. The mismatches adjacent to the CpG step introduce local structural flexibility in these
duplexes. Similar structural properties are also expected in the stem of the hairpins formed by (GGC) n strands (Mariappan et al., 1996).

The length and the nature of the repeat sequences influence the stability of the hairpin and its equilibrium with other higher order structures like homoduplexes, triplexes, quadruplexes, etc. (Dickerson, 1991; Catasti et al., 1994; Wemmer et al., 1985). However, hairpins are the simplest of the slippage structures capable of causing expansion during replication (Sinden and Wells, 1992). Researchers hypothesize that abnormal repeat expansions associated with GC-rich DNA triplets in human diseases (Caskey et al., 1992; Richards and Sutherland, 1994; Mahadevan et al., 1992; Verkerk et al., 1991) is because of the inherent ability of these sequences to form hairpin structures. Even though the existence of either the hairpin or tetraplex structures is yet to be established under *in vivo* conditions, the hairpin model (Mariappan et al., 1995) appears to be more reliable than the tetraplex structure in explaining the biological phenomena associated with fragile X syndrome (Mariappan et al., 1996).

NMR and UV optical spectroscopy studies of d(CTG)6, d(CAG)n and d(CTG).d(CAG)n (n = 2 and 3) studies have revealed that single stranded (CTG) n (n > 2) forms stable, antiparallel helical duplexes, while the single stranded (CAG) n required at least three repeating units to form a duplex (Smith et al., 1995). NMR and UV melting experiments have shown that the melting temperature increases in the order of [(CAG) 3]2 < [(CTG) 3] 2 << (CAG) 3. (CTG) 3. The (CTG) 3 duplex is stable and exhibits NMR spectral characteristics similar to those of B-form DNA that are stabilised by Watson-Crick G-C base pairs and hydrogen bonded T.T mismatches (Smith et al., 1995). This order of relative stability of the CAG and CTG repeating sequences is also observed by
electrophoretic measurements of the longer (CAG)$_{15}$ and (CTG)$_{15}$ sequences (Mitas et al., 1995). Thus, the presence of A.A mismatches in the sequence context of (CXG)$_n$ has more profound destabilization effect compared to that of the T.T mismatch. In addition to these results, recent in vitro and primer-extension studies have also suggested that longer CTG repeats might form hairpin structures, which are capable of causing interruptions in replication, leading to dynamic expansion or deletion of triplet repeat numbers (Smith et al., 1995).

It has been suggested that gene expansion occurs as a result of hairpin formation of long stretches of these trinucleotide repeat sequences on the leading daughter strand synthesized during DNA replication (Gellibolian et al., 1997). The biophysical basis of this proposed model was tested by researchers from Sheardy’s group by circular dichroism (CD) spectropolarimetry, optical melting studies, and differential scanning calorimetry (DSC) studies. Their basic objective was to evaluate the influence of sequence context and oligomer length on such secondary structures and their stabilities. After conducting these studies, it was clear that even single oligomers, as short as 12 nucleotides could also form stable hairpin structures at 25 °C. Such hairpins were characterized by the presence of N:N mismatched base pairs sandwiched between G: C base pairs in the stems and loops of three to four unpaired bases. It was then proposed that such hairpin structures had a stem region that was similar to the B-form of DNA. They were then thought to possess a B-like DNA structure, where every third mismatched nucleotide introduced a “bulge” or a localised distortion to the DNA backbone. Such distortions made them less stable than fully base-paired duplexes of the same length. Electrophoretic mobility of repeat carrying solitary sequences, as single
bands on a native gel indicated the formation of unimolecular structures that were suggested to be hairpins. CD spectra of such oligomers containing CAG, CTG and CCG repeats were strikingly similar to the CD spectra of B-form DNA. The CD spectra and native gel electrophoresis of duplexes formed by complementary strands of oligomers of these trinucleotide repeats, revealed similar spectral characteristics and gel mobility pattern. This again suggested that such complementary strands might have annealed to form B-form of DNA duplex structures (Paiva and Sheardy, 2004). Free energies of the calorimetric analysis of such probable hairpin structures transitioning into single strands showed that their stabilities were influenced by both the sequence and length of the concerned oligomer. Specifically, the stability order observed was in the order of CGG > CTG > CAG > CCG (Paiva and Sheardy, 2004). Their stability plateaued above 45 nucleotides, where the stable Watson-Crick G-C base pairs could no more compensate for the instability of the N:N mismatches in the hairpin stems (Paiva and Sheardy, 2004).

Purine-purine mismatches under different solutions have been shown to base pair in various configurations (Lane and Peck, 1995). Both x-ray crystallography and NMR studies have confirmed that G.G mismatches could adopt G\textsubscript{(anti)}-G\textsubscript{(syn)} mispairs conformation with a very different hydrogen bonding pattern (Skelly et al., 1993). At the same time, NMR studies have also suggested the presence of weakly hydrogen-bonded G\textsubscript{(anti)}-G\textsubscript{(anti)} pairs in solution (Borden et al., 1992). The pairing of mismatched G residues is highly dynamic and suggests the flipping of G between syn and anti conformations about the glycosidic bond. Hence, G.G base pairs destabilise the B form of DNA helix (Lane and Peck, 1995).
Thermodynamic measurements were reported for DNA duplexes with internal A.A, C.C, G.G, and T.T single mismatches in all the possible Watson-Crick nearest-neighbouring nucleotides’ context. The observed trend in stabilities of mismatches at 37°C was G.G > T.T ≈ A.A > C.C. The stability order for the closing Watson-Crick pair on the 5’ side of the mismatch was G.C ≥ C.G ≥ A.T ≥T.A. In addition, NMR spectra also showed that G.G and T.T mismatches were forming hydrogen-bonded structures that could vary depending on the Watson-Crick context (Peyret et al., 1999).

Salt-dependent differential scanning calorimetric measurements of oligomers having various triplets placed one after the other without being repeated and containing either GC, AT, GG, CC, AA, or TT pairs positioned in the middle, threw new light on thermodynamic stability. With reference to the central AXT/ TXA triplet (X being the position of mismatched nucleotide), the thermal and thermodynamic stabilities of the duplexes change in the following order: GC > AT > GG > AA ≈ TT > CC. The thermodynamic impact of the GG, AA, and TT mismatches were confined within the central triplet while that of the CC mismatched pair gets propagated into the adjacent helix domains and could involve 7–9 base pairs (Tikhomirova et al., 2006).

To sum up all the major structural probing studies on the triplet repeats, it could be said that such repeat sequences are capable of forming secondary structures of hairpins, triple helices and tetraplexes, which are more stable when compared to a B-form double helix. The association of disease causing triplet repeat sequences with that of the presence of unusual stable secondary structures, points towards the significance of such repeats towards the mechanism of repeat instability. During the cellular functioning of DNA replication, transcription and repair, single stranded DNA containing repeat
sequences are exposed. This allows the triplets to form the stable secondary structures (Pearson and Sinden, 1996; Gacy et al., 1995). These repetitive triplet sequences and aberrant secondary structures put together can be thought to be somehow responsible in failing the DNA repair mechanism which then can no longer change the unusual aberrant structures back to regular B-form DNA. Despite of all the intensive research, there are still many questions that need to be answered. It is absolutely necessary to get a clear understanding of the reason behind the association of diseases with repeat number instability. Questions like the molecular mechanisms leading towards instability, bias only towards expansion over time, reasons for secondary structure formation etc. definitely needs our attention. They leave the field open for more scope of structural probing studies. This is where DNA charge flow quenching mapping could prove to be useful as a tool to learn more about the unusual secondary structures formed by the repeats.

DNA charge flow through double stranded helices of Watson-Crick base pairs can conduct charges for a considerable distance of 200 Å. The extent of stacking of the DNA bases is responsible for the degree of charge flow through it. Perturbations in the form of mismatched bases, bulged bases could affect this stacking which in turn greatly affects the degree of charge flow. In chapter two, we have already studied the effect of introduction of mismatched bases in a DNA helix on the behaviour of charge flow.

Here we are using duplex DNA to model charge transfer in CNG repeat sequences. DNA strands containing triplet repeat sequences of the CNG form (N being any of the four DNA oligonucleotides) are palindromic. Thus a single strand of repeat-containing DNA is self complementary and can form duplexes with mismatches at each
N. Barring the positions of the N-N mismatched bases, all other cytosine and guanine nucleotides could be Watson-Crick base paired. It was therefore hypothesised that DNA strands comprising repeats of the type of CNG could also base pair with their complementary GNC repeats, resulting in the formation of a secondary structure that might closely resemble a double helix. The helix could be imagined as having single N:N mismatched bases sandwiched between two Watson-Crick G-C pairs. We were interested in examining the non-canonical mismatched pairs of A.A, T.T, G.G and C.C, since not much information is available about them that would help in understanding the structural features of CNG repeats. The presence of bases stacked inside a double helix is the major requirement for charge transfer. This system should provide insight into the structure of these repeats.
3.2 Materials and methods

3.2.1 Synthesis of NHS ester of Anthraquinone-2-carboxylic Acid

To label DNA with 5’ amino modification with anthraquinone, a derivative of anthraquinone with carboxylic acid, (namely anthraquinone-2-carboxylic acid) was directly coupled to a 5’ amino group on the end of a C6 linker of the DNA. The coupling was done by the formation of a peptide bond, which ultimately led to the formation of anthraquinone labelled DNA (Fahlman, 2001).

3.2.2 DNA Synthesis and Purification

The DNA oligonucleotides were synthesized at the University of Calgary Core DNA Services. The oligonucleotides were then treated following the same procedure as mentioned in section 2.2.2.

3.2.3 C6-Amino DNA Synthesis and Purification

The DNA oligonucleotides were synthesized at the University of Calgary Core DNA Services. The oligonucleotide to be derivatized with AQ was synthesized with a 5’-C6-amino functionality. Upon receipt, the DNA was treated by following the same procedure as mentioned in section 2.2.3

3.2.4 Anthraquinone Coupling of C6-Amino DNA

With the purified stock of C6-amino DNA, AQ coupling reaction was carried out. The procedure of AQ coupling was same as that mentioned in section 2.2.4.
3.2.5 HPLC Purification of Anthraquinone-coupled DNA

The AQ-coupled DNA was purified by high-pressure liquid chromatography (Agilent 1100 series HPLC) using a reverse-phase C-18 column. The solvent flow was continuously 1 ml/minute and the solvent was heated to 40 °C. The initial conditions were 100% solvent A (20:1 50 mM TEAA (pH 6.85): acetonitrile), changing to 40% solvent B (100% acetonitrile) over 30 minutes with a linear gradient. Following this period, the solvent was changed to 100% solvent B for 10 minutes to flush out and clean the column before reconditioning the column to the original conditions. Typically, uncoupled DNA has a retention time of approximately 11 minutes, while the AQ-DNA has a retention time of approximately 15 minutes. Retention times vary for different sequences depending on the length of the DNA. To ensure that the collected fractions were indeed AQ-DNA a ratio of the area of the absorbance peak at 260 nm was divided by the area of the absorbance peak at 335 nm was calculated. AQ-DNA should have a ratio range (100-120) depending on the length and sequence of the DNA (Fahlman and Sen, 2002). The fractions containing the AQ-DNA were lyophilized under vacuum and resuspended in 50 μl ddH₂O. The AQ-DNA concentration was determined by UV absorption at 260 nm using a Varian Cary UV/Vis spectrophotometer. A typical yield of AQ-DNA conjugates ranges from 50- 80% depending on the sequence and synthetic batch. The exact retention times of the DNA oligonucleotides will vary depending on its length, but in all cases, the uncoupled DNA will elute from the reverse-phase HPLC column before the coupled DNA. Absorbance spectra were taken at 260 nm (A) and 335 nm (B). DNA absorbs only at 260 nm and AQ absorbs at both 260 and 335 nm, therefore the peak at approximately 11 minutes is the uncoupled DNA and the peak at
approximately 15 minutes is the DNA coupled to AQ. The final broad peak at 22 minutes is the uncoupled AQ.

### 3.2.6 5’ End-Labelling

DNA oligonucleotides were 5’ end-labelled using a standard phosphorylating procedure (see appendix for procedure). The DNA was resuspended in 20 μl denaturing gel loading buffer and purified on 10% (w/v) denaturing polyacrylamide gel. The purified DNA was resuspended in 30 μl TE (10 mM Tris-Cl pH 7.5 and 0.1 mM EDTA pH 8.0).

### 3.2.7 Assembly of DNA Constructs

DNA assemblies were formed by annealing stochiometric mixtures of two constituent oligonucleotides (1 μM each), one of which is either the AQ-DNA or C6 amino linked uncoupled- DNA and also small amounts of 5’-end labelled oligonucleotide (that was sufficient to serve as a tracer), in 10 mM Tris-Cl (pH 7.5) and was made up to a volume of 20 μl. The solutions were heated in a boiling water bath at 100 °C for 2 minutes, cooled slowly to room temperature (about 60 minutes) in a PCR heat exchanger. Then it is diluted 2-fold in irradiation buffer (final composition: 5 mM Tris-Cl (pH 7.5), 95 mM LiCl, 1 mM MgCl₂) and made up to a volume of 40 μl. The samples were covered and incubated for approximately 15 minutes at room temperature prior to photo-irradiation to initiate charge flow.

### 3.2.8 Photo-irradiation of the DNA duplex assembly

DNA solutions after being annealed were placed in the PCR heat exchanger set to different temperatures as required (21 °C normally is taken as room temperature, 10 °C) for irradiation. A Black-Ray UVL-56 lamp (365 nm) is kept just above these
microcentrifuge tubes for 60 minutes for irradiation. For some temperature-dependent charge transfer experiments, additional time periods of 30 minutes, 60 minutes, 120 minutes and 180 minutes photo-irradiation were also used. Following photo-irradiation, the samples were ethanol precipitated with the assistance of glycogen.

3.2.9 Dimethyl Sulphate Modification of DNA (Guanine Ladder)

Maxam-Gilbert chemical DNA sequencing Dimethyl sulphate (DMS) ladder (see appendix for complete procedure). DMS treatment modifies guanine bases by addition of a methyl group. After DMS modification, the DNA pellets were dissolved in 90 µl of ddH₂O and 10 µl of fresh piperidine was added for piperidine cleavage. The solution was incubated at 90 °C for 30 minutes in the dark and lyophilised.

3.2.10 Piperidine Treatment and Denaturing Gel Electrophoresis

The DNA pellets were dissolved in 100 µl 10% (v/v) piperidine and incubated at 90 °C for 30 minutes. The treated DNA was lyophilized, dissolved in denaturing gel loading dye (99% formamide (v/v), 10 mM Tris-Cl pH 7.5, and 0.1mM EDTA pH 8.0), heat denatured at 100 °C for 5 minutes, cooled to 21°C by putting the samples on ice. The samples were then loaded on 15% (w/v) denaturing polyacrylamide sequencing gels. The gels were analyzed by phosphorimagery using a Molecular Dynamics Typhoon 9410 Variable Mode Imager.

3.2.11 Data Analyses

The densitometry analyses of the sequencing gels were carried out using Molecular Dynamics Image Quant 5.2 software for Windows 2000. The density of the band of interest was assessed as a percentage of the total signal contained within a
particular lane. This procedure compensated for any discrepancies of the total radioactive counts in different lanes. The normalized signal for a particular nucleotide obtained in the "dark" reaction negative control lane was subtracted from the signal of the same nucleotide in an irradiated sample lane.
3.3 Results and Discussion

3.3.1 Experimental Design of DNA duplexes containing trinucleotide repeats of the CNG type (N= Adenine/ Guanine/ Cytosine/ Thymine)

The application of charge transfer in the field of structural probing has been used to study complex DNA assemblages like those of three-way junctions, four-way junctions; multi-helix DNA constructs involved in aptamer-ligand binding; folded assemblies such as the pseudosubstrate bound 8-17 deoxyribozyme (Santhosh and Schuster, 2003; Odom et al., 2001; Fahlman et al., 2002; Sankar and Sen, 2004; Fahlman and Sen, 2002; Leung and Sen, 2007). Our laboratory has also been interested in looking into DNA and RNA structures and has its own share of contribution towards such structural studies using charge flow as the tool (Fahlman and Sen, 2002; Leung and Sen, 2007; Bergeron and Sen, 2008).

As mentioned earlier, several biophysical studies have shown that TNRs of the CNG type, form alternative DNA secondary structures \textit{in vitro} with two repeating G-C Watson–Crick pairs and a mismatched N:N pair, sandwiched between them. DNA duplexes carrying such repeating trinucleotides were to be tested with UV light initiated DNA charge transfer experiments.

Four DNA oligomers (43 nucleotides in length), each of which had three sets of either of the CAG/CGG/CCG/CTG repeats were designed. The three identical repeats were introduced sequentially, starting from the 18\textsuperscript{th} position of the radio-labelled 5’-\textsuperscript{32}P-end of the oligomer. This placed the repeats in the mid region and continued with two 17 nucleotide arms on each side of them. Both the 17 nucleotide arms comprised of the base paired Watson-Crick region for the DNA oligomers. For the duplex formation, four more
DNA oligomers were designed that were complementary to the former four DNA oligomers excluding the repeat region. The complementary strands had the identical triplet repeats placed in the same sequence and position and a $\gamma^{32}\text{P}$ labelled end. Hence three mismatched N:N pairs were introduced into individual sets of duplexes. For all the four mismatched duplexes, control duplexes were also designed that were completely Watson-Crick base paired. The DNA oligomers were designed such that no two G-C pairs are distanced from each other more than one A-T base pair apart. This ensures that the A-T base pairs that are known to inhibit charge flow; do not hinder the transfer through the primary charge carriers of G-C base pairs (Fig 3.1, Fig 3.2). The 5’-end of AQ coupled strand was the point of injection of charge. The sequence at this end was chosen to be 5’TTTA3’. This ensured better propagation of the charge in the forward direction than in the backward direction, so as to prevent charge recombination. Since the A-T base pairs inhibit charge flow, so once the charge is introduced at the closest guanine, the possibility of the charge going backward to recombine with AQ$^{\cdot}$ would be diminished.

Charge was injected from the 5’ end of the AQ labelled strand by the covalently tethered anthraquinone molecule. The 5’ end of the complementary strand was radio-labelled with $\gamma^{32}\text{P}$. Two sets of guanine doublets were placed, one on each side of the triplets to act as charge traps. Two GG doublets positioned at 13th, 14th and 30th, 31st positions from the 5’ end of the labelled DNA strand were referred as distal and proximal doublets respectively. The N:N mismatched pairs were at positions 19, 22 and 25 from the 5’ labelled strand (Fig.3.1).
All densitometry signals were calculated as a percent based on the counts of the band of interest divided by the total counts of that lane. The signal for charge transfer dependent oxidative damage of a particular guanine base was obtained by subtracting the background percent damage of the same corresponding base from the control reaction which was missing the tethered anthraquinone. All “proximal:distal damage ratios” were calculated by dividing the corrected percent damage of the 5’ guanine of the distal doublet (G13) by the 5’ guanine of the proximal doublet (G30). The “negative control” for all the experiments in this chapter consisted of UV light irradiated DNA duplexes containing C6-amino linkers that were not coupled to AQ. The background reaction shown by these negative controls was not guanine specific and was caused due to the usual damaging action on the DNA, exposed to UV light. They were believed to serve as better controls as compared to the negative controls consisting of AQ duplexes that were not exposed to UV light (as in the experiments of chapter 2).

The DNA duplexes were assembled with their respective pairs of single strands in the presence of salts. The assembled duplexes were run on non-denaturing native gels along with 100 b.p. ladders. Single bands on the gels showed that no single strands were present and that duplexes are more likely to be formed than other single strand structures. This was done prior to the start of any charge transfer experiment (data not shown).
Figure 3.1  Constructs for DNA duplex carrying mismatched CNG triplet repeats, where the N:N mismatched pairs are G.G, A.A, T.T and C.C respectively. Bases in red denotes the triplet repeat region, bases in black are Watson-Crick base paired. The DNA strands whose repeats are to be studied are 5' 32P labelled. AQ is derivatized on the 5' end of the DNA strand opposite to the radio-labelled strand. The blue thick lines denote the positions, where the mismatched N:N pairs are present.
CCG repeat (has G-C base pairs)

TTTAGCTCAGCTCTCTCCTCGCGCGCGGACTCCATAGTGAGTAAG
AAATCGAGTCGAGGAGAAAGCGCGCGCGCCGTGAGGTATCCTCATTCP\textsuperscript{32}5',

CGG repeat (has C-G base pairs)

TTTAGCTCAGCTCTCTCCTCGCGCGCGGACTCCATAGTGAGTAAG
AAATCGAGTCGAGGAGAAAGCGCGCGCGCCGTGAGGTATCCTCATTCP\textsuperscript{32}5',

CTG repeat (has A-T base pairs)

TTTAGCTCAGCTCTCTCCTCGCGCGCGGACTCCATAGTGAGTAAG
AAATCGAGTCGAGGAGAAAGCGCGCGCGCCGTGAGGTATCCTCATTCP\textsuperscript{32}5',

CAG repeat (has T-A base pairs)

TTTAGCTCAGCTCTCTCCTCGCGCGCGGACTCCATAGTGAGTAAG
AAATCGAGTCGAGGAGAAAGCGCGCGCGCCGTGAGGTATCCTCATTCP\textsuperscript{32}5',
Figure 3.2  Constructs for Watson- Crick control DNA duplex carrying CNG triplet repeats, where ‘N’ is cytosine, guanine, thymine and adenine respectively (all the ‘N’ bases are Watson –Crick base paired with their respective complementary bases). Bases in red denotes the triplet repeat region, bases in black are Watson-Crick base paired. The DNA strands whose repeats are to be studied are 5’ $^{32}$P labelled. AQ is derivatized on the 5’ end of the DNA strand opposite to the radio-labelled strand. Blue thick lines denote the hydrogen bonding between Watson-Crick A-T base pairs.
3.3.2 Could DNA duplex carrying CCG, CGG, CTG and CAG repeats conduct charge?

Figure 3.3 shows the figure of a denaturing gel showing charge flow-dependent DNA strand cleavage patterns in four mismatched duplexes carrying three copies of CCG, CGG, CTG and CAT repeats each. Thus each of these CNG repeats carrying duplexes (N= C, G, T, A), individually has three C.C, G.G, T.T and A.A mismatched pairs respectively. There are also four Watson-Crick duplex controls and four Maxim-Gilbert sequencing G ladders present along with their respective group of mismatched duplexes. Lanes 5, 10, 15, 20 shows the piperidine cleavage product of photo-irradiated AQ-derivatized DNA mismatched duplexes carrying their respective N:N mismatched pairs. Lanes 4, 9, 14 and 19 shows the piperidine cleavage product of photo-irradiated mismatched duplexes having C6-amino linker, but are not tethered to anthraquinone. Lanes 3, 8, 13, 18 demonstrate the piperidine cleavage product of photo-irradiated AQ-derivatized Watson-Crick control duplexes carrying G-C, C-G, A-T and T-A W.C. base pairs respectively. Lanes 2, 7, 12, 17 shows the piperidine cleavage product of photo-irradiated Watson-Crick control duplexes having C6 amino linker, but is not tethered to anthraquinone. Lanes 1, 6, 11, 16 demonstrate Maxam-Gilbert Dimethyl sulphate (DMS) sequencing G ladders, running along with their respective set of mismatched and control duplexes.

All the duplexes have a GG doublet strategically placed right before (G30, G31 – proximal G doublets) and after the start and end (G13, G14 – distal G doublet) of the repeat region. On comparing the ratios of charge flow between these GG doublets in the mismatch duplex with respect to their Watson-Crick duplex controls, we can try to
comment on the efficiency of charge transfer across the mismatched pairs. This in turn might give an insight on the stacking ability of these mismatches present between Watson-Crick base pairs.

There was considerable charge flow through all four of the mismatched TNR regions of the duplex (Figure 3.3). For comparison of conductivity, every set of mismatch TNR duplexes had a corresponding Watson-Crick control duplex as a positive control.

Densitometry of the signal of C.C, G.G, T.T and A.A carrying mismatched duplexes (lane 5, 10, 15 and 20 of Figure 3.3) were measured to be approximately 9%, 8%, 12% and 11% of the total of their respective lanes; whereas the counts of their corresponding Watson-Crick control duplex (lane 3, 8, 13 and 18 of Figure 3.3) was also approximately around 9% 15%, 16% and 16% of the total of the respective lanes. The values presented here are based on the result of six independent sets of experimental data sets having a standard deviation value between 2% and 3%. This shows that there does not exist a dramatic difference in charge flow-dependent damage between the mismatched duplexes carrying three of C.C, G.G, T.T and A.A mismatched pairs and their corresponding Watson-Crick control duplexes (Figure 3.4).

From chapter 2 (Figure 2.3.6) we reported that introduction of a single mismatch pair in a DNA duplex could reduce the efficiency of charge flow, if either of the mismatched pair is not a guanine. Thus it would not be absolutely unexpected that single G.G mismatched pairs could possibly share the similar range of oxidative-damage, when compared to its Watson-Crick control. Barton’s group in 2001 had also showed that Watson-Crick G-C and G.G mismatched pairs exhibited similar damage supported by
charge flow (Bhattacharya and Barton, 2001). This would possibly account for the damage seen (in Figure 3.3) through G.G mismatched pairs.

However, A.A mismatched pairs in the same duplex could be expected to interfere with the charge flow and reduce the damage that is seen on the gel (Figure 2.3.6). Mismatched pairs of the kinds of C.C and T.T pairs were thought to be significantly interfering with the charge flow and hence reducing the damage levels to an insignificant level. However, the observation from Figure 3.3 shows that all the four duplexes carrying C.C, G.G, T.T and A.A mismatched pairs were capable of producing charge flow related damage, which was comparable to their Watson-Crick controls. We know that guanine is the primary charge carrier where most of the damage is seen. All four duplexes had guanine rich CNG repeats. The presence of many guanines in either of the strands in the repeat region may be responsible for such an observation. In spite of having three single mismatched pairs in the duplex, the cleavage observed was not markedly different from that of the Watson-Crick control duplexes. Hence, our hypothesis that the introduction of single mismatched pairs in a DNA duplex would make a substantial impact on the oxidative damage, is proved to be incorrect. Our results show that all A.A, G.G, T.T and C.C mismatched pairs, conducted charges at approximately the same efficiency as that of the Watson-Crick base pairs. This is most likely due to the strong charge flow characteristics of the adjacent guanine residues.
Figure 3.3 Denaturing gel showing charge flow-dependent DNA strand cleavage patterns in mismatched duplexes along with their corresponding Watson-Crick duplex controls. Lanes 5, 10, 15, 20: Photo-irradiated AQ-derivatized (with C6 amino linked DNA) duplexes carrying C.C, G.G, T.T and A.A mismatched pairs respectively. Lanes 4, 9, 14, 19: Photo-irradiated C6 amino linked duplexes carrying C.C, G.G, T.T and A.A mismatched pairs respectively. Lanes 3, 8, 13, 18: Photo-irradiated AQ-derivatized (with C6 amino linked DNA) duplexes carrying G-C, C-G, A-T and T-A Watson-Crick base pairs respectively. Lanes 2, 7, 12, 17: Photo-irradiated C6 amino linked duplexes carrying G-C, C-G, A-T and T-A Watson-Crick base pairs respectively. Lanes 1, 6, 11, 16: Maxam-Gilbert Dimethyl sulphate (DMS) G ladders, running along with their respective group of, mismatched and their corresponding control triplet repeat duplexes.

![Total Cleavage %](image)

Figure 3.4 Total cleavage percentage for duplexes containing C.C, G.G, T.T and A.A mismatched pairs along with their corresponding Watson-Crick controls respectively.
3.3.3 How would any change in temperature during photo-irradiation affect the charge flow through DNA duplex?

Charge transfer is sensitive to changes in temperature that might occur during irradiation. In our laboratory, 21 °C is the usual room temperature at which all my charge transfer experiments are carried out. If irradiation is done at a lower temperature of 10 °C, we might expect changes in the pattern of oxidative damage shown in the gel. The stacking interactions in a helix are never static but are modulated by structural dynamics (O’Neill and Barton, 2002). In a system of DNA solution, structural dynamics is the spontaneous fluctuation of its helical structure into different conformational states that are energetically similar to each other (O’Neill and Barton, 2002). Thus the bases in a duplex are always in a state of conformational motion. Any change in the temperature is expected to change the energy and dynamics of the experimental DNA system. This in turn leads to changes in the helical stacking of the duplex, which is the most crucial parameter involved during charge flow. Changes in temperature that induces conformational changes, might be responsible for charge flow through such altered conformations (O’Neill and Barton, 2002). It could ultimately be expected to affect the conductivity pattern and oxidative damage (Núñez et al., 1998).

To determine whether any change in temperature during photo-irradiation affects charge flow, we selected one of the four DNA duplex, containing CTG repeats (with T.T mismatched pairs) and its corresponding Watson-Crick control duplex (Figure 3.5). It is already known that guanines act as good carriers of charge; hence a probability of higher oxidative damage might be expected in spite of having three single mismatched bases. The T.T mismatch is shown to strongly inhibit charge flow and thus the CTG repeat was used for these experiments. Figure 3.6 is the image of a denaturing gel showing charge
flow-dependent DNA strand cleavage patterns of CTG repeat containing Watson-Crick duplex control (having A-T base pairs) at two different temperatures of 10 °C and 21 °C with irradiation time periods of 30, 60, 120 and 180 minutes.

Figure 3.7 is the image of a denaturing gel showing charge flow-dependent DNA strand cleavage patterns of CTG repeat containing T.T mismatched duplex (carrying three T.T mismatched pairs) at two different temperatures of 10 °C and 21 °C with irradiation time periods of 30, 60, 120 and 180 minutes. All of the duplexes have a GG doublet strategically placed right before (G30, G31 – proximal G doublets) and after the start and end (G13, G14 – distal G doublet) of the repeat region.

The mismatched CTG-TNR duplex showed higher intensity of damage than the Watson-Crick CAG control duplex at both 21 °C and 10 °C. The conductivity might remain the same as it was before the temperature change, yet the damage might increase because of greater exposure of the bases to the solvent. For a duplex undergoing 60 min irradiation, temperature does not have much effect on the damage. With increase of irradiation time (120 min), the duplex shows greater damage at both 21 °C and 10 °C. At 120 min irradiation, the difference in the damage between a duplex and its Watson-Crick CAG control is highest, both at 21 °C and 10 °C.

Densitometry of the signal of both the T.T. mismatched and its Watson-Crick control duplexes were calculated at both 10 °C and 21 °C for all the irradiation time periods of 0, 30, 60, 120 and 180 minutes. Experiments were performed four times independently, to obtain four data sets. The average of all of them was supposed to be taken into account, but the standard deviation values were quite high as compared to the values of my data sets making them statistically insignificant to be reported.
We can at the most, assume that 60 minutes probably could be chosen as the irradiation time period at both 21 °C and 10 °C temperatures. At any higher temperatures, the damage might go beyond the limit of linear nature of charge transfer. Though the oxidative damage in mismatched duplex does increase with increasing temperatures, the increase was not statistically significant. This unfortunately leaves us with no definite conclusion. It would be difficult to comment on such a conductivity pattern since our past experiences with charge transfer experiments have suggested that the technique itself is capable of creating variations in results that becomes statistically significant with respect to the small value of total cleavage obtained from the experiments.
Figure 3.5 Construct for DNA duplex carrying mismatched CNG triplet repeats, where the N:N mismatched pair is T.T. Construct for its corresponding Watson-Crick control DNA duplex, where ‘N’ is adenine (adenine is Watson–Crick base paired with its complementary thymine base). Bases in red denote the triplet repeat region, bases in black are Watson-Crick base paired. The DNA strands whose repeats are to be studied are 5’ $^{32}$P labelled. AQ is derivatized on the 5’ end of the DNA strand opposite to the radio-labelled strand. Black circles denote the positions where T: T mismatches are present. Blue thick lines denote the hydrogen bonding between Watson-Crick A-T base pairs.
Figure 3.6  Denaturing gel showing charge flow-dependent DNA strand cleavage patterns of CTG repeat containing Watson-Crick duplex control (having A-T base pairs) at different temperatures as a function of various photo-irradiation time periods. Lanes 2, 4, 6, 8 and 10: Photo-irradiation of AQ-derivatized (with C6 amino linked DNA) W. C. duplex at 10 °C, for the time periods of 0, 30, 60, 120 and 180 minutes respectively. Lanes 1, 3, 5, 7 and 9: Photo-irradiation of C6 amino linked W. C. duplex at 10 °C, for the time periods of 0, 30, 60, 120 and 180 minutes respectively. Lanes 12, 14, 16, 18 and 20: Photo-irradiation of AQ-derivatized (with C6 amino linked DNA) W. C. duplex at 21 °C, for the time periods of 0, 30, 60, 120 and 180 minutes respectively. Lanes 11, 13, 15, 17 and 19: Photo-irradiation of C6 amino linked W. C. duplex at 21 °C, for the time periods of 0, 30, 60, 120 and 180 minutes respectively. Lane 21: Maxam-Gilbert Dimethyl sulphate (DMS) G ladder for the Watson-Crick control duplex.
Figure 3.7 Denaturing gel showing charge flow-dependent DNA strand cleavage patterns of CTG repeat containing T.T mismatched duplex at different temperatures as a function of various photo-irradiation time periods. Lanes 3, 5, 7, 9 and 11: Photo-irradiation of AQ-derivatized (with C6 amino linked DNA) T.T mismatched duplex at 10 o C, for the time periods of 0, 30, 60, 120 and 180 minutes respectively. Lanes 2, 4, 6, 8 and 10: Photo-irradiation of C6 amino linked T.T mismatched duplex at 10 o C, for the time periods of 0, 30, 60, 120 and 180 minutes respectively. Lanes 13, 15, 17, 19 and 21: Photo-irradiation of AQ-derivatized (with C6 amino linked DNA) T.T mismatched duplex at 21 o C, for the time periods of 0, 30, 60, 120 and 180 minutes respectively. Lanes 12, 14, 16, 18 and 20: Photo-irradiation of C6 amino linked T.T mismatched duplex at 21 o C, for the time periods of 0, 30, 60, 120 and 180 minutes respectively. Lane 1 and 22: Maxam-Gilbert Dimethyl sulphate (DMS) G ladders for the T.T mismatched duplex.
a) 10, 21 °C: Total cleavage at 30 min irradiation

<table>
<thead>
<tr>
<th></th>
<th>10D W.C.</th>
<th>10D Mis.</th>
<th>21D W.C.</th>
<th>21D Mis.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cleavage</td>
<td>3.77</td>
<td>5.82</td>
<td>4.3247</td>
<td>7.97</td>
</tr>
<tr>
<td>% Cleavage at 30' irradiation</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

W.C. and Mismatched duplex at 10, 21 °C

b) 10, 21 °C: Total cleavage at 60 min irradiation

<table>
<thead>
<tr>
<th></th>
<th>10D W.C.</th>
<th>10D Mis.</th>
<th>21D W.C.</th>
<th>21D Mis.</th>
</tr>
</thead>
<tbody>
<tr>
<td>% Cleavage at 60 min irradiation</td>
<td>7.85</td>
<td>10.62</td>
<td>7.95</td>
<td>10.24</td>
</tr>
</tbody>
</table>

W.C. and Mismatched duplex at 10, 21 °C
Figure 3.8  Total cleavage percentage for CTG repeat containing mismatched and Watson-Crick duplex at 10 °C and 21 °C for the time period of (i) 30 min (ii) 60 min (iii) 120 min (iv) 180 min. (10 D = 10 °C, 21 D = 21 °C; W.C. = Watson-Crick control duplex, Mis. = T.T Mismatched duplex)
3.4 Conclusion

We have seen in the previous chapter that the studies of the oxidative damage pattern of charge flow through DNA and RNA molecules patterns is able to provide information about stacking ability of nucleic acid duplexes containing single mismatches. The extent of base stacking between adjacent bases or base pairs was thought to influence the conductivity pattern. DNA duplexes carrying three repeats of each of CAG, CGG, CTG and CCG sequences in the middle region, have all shown considerable charge flow through their mismatched repeat regions. When compared to their respective positive control duplexes (in which all the bases are paired), the difference in the oxidative damage between the control and mismatched duplexes is not statistically significant.

Various time periods were used to observe conductivity and all of the DNA duplexes carrying mismatched base pairs showed saturation beyond two hours of sample irradiation, after which there was no further increase in damage with the increase in irradiation time period. Two different temperatures (10 °C and 21 °C), combined with all the different time periods were used to conduct the charge flow experiments. Once again, though the oxidative damage increased with increasing temperatures but the increase was not statistically significant.
3.5 Future Directions

We have seen that charge flow through DNA duplexes carrying three copies of a triplet repeat does not give promising results to continue with the same DNA constructs. Introducing higher number of repeats could be one of the alternatives. New DNA constructs could be designed that would carry four or more copies of the triplets placed consecutively in the DNA strands. This would increase the total number of mismatched pairs that the charge has to pass before it could hit the distal GG doublet charge trap. We might expect that the oxidative damage pattern of the mismatch containing duplexes would be significantly different from those of their corresponding Watson-Crick control duplexes.
Chapter 4: Final Conclusion

DNA-DNA duplexes with guanine-rich termini generate a large oxidative end-effect. Comparative charge flow experiments using piperidine and aniline exhibited the end-effect. Introduction of certain single-base mismatches within the duplex decreased the magnitude of this end-effect. Hence, this intense magnitude of end effect has the potential to act as a highly sensitive reporter of charge flow in DNA duplexes. RNA-RNA duplexes also demonstrate the same massive oxidative end-effect. To our knowledge, it was the first demonstration of charge flow through a RNA-RNA double helix. We can conclude that both the A and B family of nucleic acids do conduct charges comparably and also have the potential to act as sensitive reporters of charge flow. This property could prove to be useful for studying the structure, folding, and dynamics of complexly folded RNAs and DNAs.

DNA duplexes carrying trinucleotide repeats of CAG, CGG, CTG and CCG sequences have all shown considerable charge flow dependent damage. In spite of having three single mismatches separated from each other by two G-C Watson-Crick base pairs, charge flow is not much interrupted. When compared to their respective positive control duplexes (Watson-Crick base paired duplex), difference in the oxidative damage between the control and mismatched duplexes is not statistically significant. Various time periods were used to observe their conductivity and all of the DNA duplexes carrying mismatched base pairs showed saturation beyond two hours of sample irradiation with
UV light. Two different temperatures (10 °C and 21°C), combined with all the different time periods were used to conduct the charge flow experiments. With the increase in temperature, the damage in both the T.T mismatched duplex and the Watson-Crick control duplex increased, but the increase was not statistically significant.
Appendix: Common Experimental Procedures

Ethanol Precipitation

All oligonucleotide samples were precipitated by adding 1/10th volume 3 M NaOAc (or NaCl to a final concentration of 300 mM) and 2.5 times the volume with anhydrous ethanol. At times, 1 μg of glycogen was added to the samples to act as a carrier during precipitation. The samples were vortexed, placed in dry ice until it was frozen, and centrifuged at 4°C at 16.1 RCF for 40 minutes. The supernatant was removed and the resulting pellet was resuspended in the desired solvent.

5' End-labeling

2.5 μM of the oligonucleotide was incubated with 20 units of T4 polynucleotide kinase and 1.25 μM $^{32}$P-ATP in a 20 μl reaction volume. The reaction was incubated at 37°C for 30 minutes and was ethanol precipitated as mentioned above. The oligonucleotide pellet was resuspended in 20 μl of denaturing gel loading buffer and purified by denaturing polyacrylamide gel electrophoresis. The oligonucleotide band was visualized by using film autoradiography and excised from the gel. The sample was eluted by crush-soak methodology using 1 ml TE overnight. The gel pieces were briefly spun down and the supernatant was removed and placed into new microcentrifuge tubes. The samples were ethanol precipitated as previously mentioned, resuspended in a final volume of 30 μl ddH$_2$O or TE, and stored at -20°C.
**Dimethyl Sulfate Modification of DNA (G Ladder)**

The modification reaction was performed using 1.25 μM DNA oligomers and trace amounts of 5'-^32^P end-labeled of the DNA oligomer in 50 mM NaOH/ cacodylate (pH 7.5). The sample was then made to a volume of 35 μl using ddH₂O, heated at 100°C for 1 minute, briefly spun in a microcentrifuge, and equilibrated at 22°C for 10 minutes. Dimethyl Sulfate (DMS) modification was initiated by adding 5 μl 4% (v/v) DMS and the reaction was incubated at 22°C for 30 minutes. The reaction was quenched by adding 10 μl DMS Stop Solution (1 M β-Mercaptoethanol and 1.5 M NaOAc) and ethanol precipitated by adding 110 μl anhydrous ethanol. The DNA pellet was re-dissolved either in 100 μl of 10% (v/v) piperidine or aniline acetate (20 μl aniline + 180 μl H₂O + 12 μl glacial acetic acid) and heated at (90°C or 60°C) for 30 or 15 minutes respectively, and lyophilized. The sample was resuspended in 40 μl denaturing gel loading dye. Figure below shows the reaction scheme of DMS modification of guanine.

Appendix Figure 1 Guanine ladder reaction scheme (Bloomfield et al., 2000).

112
Di-ethyl pyrocarbonate (DEPC) Modification of DNA (A > G Ladder)

The modification reaction was performed using 1.25 μM DNA oligomers and trace amounts of 5'-32P end-labeled of the DNA oligomer in 25 mM NaOH/ cacodylate buffer (pH 7.5). The sample was then made to a volume of 39 μl using ddH2O, and followed by addition of 1 μl of di-ethyl pyrocarbonate (DEPC). The reaction was incubated at room temperature for 30 minutes. The sample was vortexed frequently during those 30 minutes as DEPC is insoluble in water. This was followed by ethanol precipitation by addition of 100 μl anhydrous ethanol and 4 μl of 3 M sodium acetate. The DNA pellet was re-dissolved either in 100 μl of 10% (v/v) piperidine or aniline acetate (20 μl aniline + 180 μl H2O + 12 μl glacial acetic acid) and heated at (90 °C or 60 °C) for 30 or 15 minutes respectively, and lyophilized. The sample was resuspended in 40 μl denaturing gel loading dye before loading on the denaturing gel.
References


Fahlman RP. 2001. DNA Nanotech: Expanding the repertoire of DNA for the assembly of nanoscale objects and electrical devices. Theses (Institute of Molecular Biology and Biochemistry) / Simon Fraser University. Thesis (Ph.D.) - Simon Fraser University. Chapter 4: 2.1: 166-168.


