MACROSOLUTE EFFECTS ON NUCLEIC ACID INTERACTIONS

A Thesis
submitted to the Faculty of the
Graduate School of Chemistry
of Georgetown University
in partial fulfillment of the requirements for the
degree of
Doctor of Philosophy
in Chemistry

By

Sarai Obando, M.S.

Washington, DC
October 22, 2008
Copyright 2008 by Sarai Obando

All Rights Reserved
MACROSOLUTE EFFECTS ON NUCLEIC ACID INTERACTIONS

Sarai Obando, M.S.

Thesis Advisor: Steven J. Metallo, Ph.D.

ABSTRACT

The intracellular environment contains a variety of solutes that cumulatively occupy a significant volume of the cell (20-30%). The high volume occupancy generates a system which is macromolecularly crowded. This crowding, also known as the excluded volume effect, can lead to an increase in the chemical activity of solutes and influence thermodynamic and kinetic values as compared to a dilute system. Synthetic, inert cosolutes were used to provide a simplified mimic of the intracellular environment, in which the DNA structures were studied. Various macrosolutes, cosolutes of molecular weight greater than 1000, and an osmolyte, a cosolute of molecular weight less than 1000, were used to create crowded conditions. In the presence of macrosolute, differential stabilization of a complementary DNA duplex over duplexes containing a single mismatched base pair was observed. Hence, crowding effectively increased the specificity of the hybridization reaction. In contrast, a study of single bulge base
duplexes demonstrated the differential stabilization of a bulge duplex over all other duplexes when in the presence of a macrosolute. Crowding studies on unimolecular systems demonstrated marginal stabilization of stem-loop structures as compared to duplexes. However, in systems with molecularities ranging from one to four, as the molecularity of a system increased, the crowding effects also increased. For all systems studied, crowding enhanced the rate constant of hybridization. As the activation energies of the systems studied increased, the crowding effects also increased. This work contributes to our understanding of the effects that the complex intracellular environment has on biochemical interactions and processes. These results are also of potential use to bioanalytical techniques which employ nucleic acid hybridization. Additionally, these results could be used to design molecular systems, which can switch structure based on the volume occupancy of the surrounding medium.
ACKNOWLEDGEMENTS

I would like to thank Dr. Steve Metallo for his direction, assistance and guidance and for always demanding my very best. I also extend my gratitude to my graduate research group colleagues. I would especially like to thank Ms. Alison Williams, for her friendship and unwavering encouragement.

A special thanks to my student undergraduate colleagues who contributed to this work. I would like to thank Ms. Jennifer Lilian Small (C’06) for her work on the BRCA1-185 WT thermodynamics and competition studies. Thank you to Ms. Karina Vivar (C’07) for her work on the 5382-WT and 5382-C thermodynamics. I would also like to thank Mr. Christopher R. MacKay (C’08) for his work on the BRCA1-185 kinetics and his contribution of figure 2.9. Finally, I would like to thank Mr. Eric D. Nellis (C’09) for his work on the multi-branch junction thermodynamics and kinetics.

To my congregation and home away from home, Pentecostal Church of God Ebenezer, I extend my heartfelt appreciation for their prayers and understanding throughout these years.

Words cannot express my gratitude to my family who enriches my life in ways they will never fully know. I thank my husband Jose, for supporting my goals and for all the sacrifices that he has made for me. I am truly blessed to have such a wonderful and loving husband. I am forever grateful to my mother and father who never let me
give up and always believed in me. I am thankful to have parents who always lift me in
prayer and are an amazing support system. I would like to thank my brother Nelson and
my sister-in-law Cindy for always encouraging me and for loving me unconditionally. I
am also thankful to my brother Daniel for his support and for always reminding me that
laughter truly is the best medicine. Finally, to all of my friends and family, thank you for
your support.
To my family
# Table of Contents

Chapter 1 – Introduction to Nucleic Acids and Macromolecular Crowding

1.1 The Chemistry of DNA ........................................................................................................ 1

1.2 The Secondary Structure of DNA ...................................................................................... 5

1.2.1 Watson-Crick Base Pairing .......................................................................................... 5

1.2.2 The Double Helix .......................................................................................................... 5

1.2.3 Other Structural Forms of DNA ................................................................................... 9

1.3 Hydration of DNA .............................................................................................................. 10

1.4 Single Base Mutations in DNA ......................................................................................... 11

1.4.1 Mismatched Bases .................................................................................................... 11

1.4.2 Bulge Bases ................................................................................................................ 16

1.5 The Crowded Intracellular Environment ...................................................................... 17

1.6 Mimicking a Crowding Environment .......................................................................... 18

1.7 Crowding Effects - Excluded Volume Theory ............................................................... 21

1.8 Crowding Effects - The Distribution of Waters and Cosolutes ....................................... 24

1.8.1 The Local Bulk Partitioning Model (LBPM) ........................................................... 26

1.8.2 The Preferential Hydration Parameter ........................................................................ 26

1.8.3 Osmotic Stress Analysis ............................................................................................. 27
Chapter 2 - Macrosolute Effects on a Bimolecular System with Single Mismatches

2.1 Introduction ........................................................................................................... 54

2.2 Materials and Methods ....................................................................................... 60

2.2.1 Materials ......................................................................................................... 60

2.2.2 Measurements ................................................................................................. 62

2.2.3 Data Analysis .................................................................................................. 66

2.3 Results ................................................................................................................... 71

2.3.1 Macrosolute Effects on the Thermodynamics of the BRCA1-185 Duplexes 71

2.3.2 Macrosolute Effects on the Kinetics of Duplex Formation of BRCA1-185 .. 85

2.3.3 Macrosolute Effects on the WT BRCA1-185 Duplex in a Competitive Environment ................................................................. 87
3.3.2 Macrosolute Effects on the Kinetics of Hybridization of the 5382-WT and 5382-C Duplexes ................................................................................................... 125

3.4 Discussion ........................................................................................................... 128

3.4.1 Crowding Mediated Stabilization of the Bulge Duplexes ......................... 128

3.4.2 Macrosolute Enhancement of the Rate Constant of Hybridization .......... 132

3.5 Conclusions ......................................................................................................... 133

References ................................................................................................................. 135

Chapter 4 - Macrosolute Effects on Unimolecular Systems

4.1 Introduction ......................................................................................................... 138

4.2 Materials and Methods ....................................................................................... 142

4.2.1 Materials ....................................................................................................... 142

4.2.2 Measurements ............................................................................................... 143

4.2.3 Data Analysis ............................................................................................... 145

4.3 Results ................................................................................................................. 150

4.3.1 Macrosolute Effects on the Stem-loops and Duplexes .............................. 150

4.3.2 Macrosolute Effects on the Kinetics of Hybridization of the Duplexes ....... 158

4.4 Discussion ........................................................................................................... 166
### Chapter 4 - Macrosolute Effects on RNA/DNA Hybrid Structures

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.4.1 Macrosolutes Marginally Enhance the Stability of Stem-loops</td>
<td>166</td>
</tr>
<tr>
<td>4.4.2 Macrosolute Effects Correlate With Increased Activation Energy</td>
<td>167</td>
</tr>
<tr>
<td>4.5 Conclusions</td>
<td>169</td>
</tr>
</tbody>
</table>

References: 170

### Chapter 5 - Macrosolute Effects on Multi-branch DNA Junctions

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.1 Introduction</td>
<td>173</td>
</tr>
<tr>
<td>5.2 Materials and Methods</td>
<td>175</td>
</tr>
<tr>
<td>5.2.1 Materials</td>
<td>175</td>
</tr>
<tr>
<td>5.2.2 Measurements</td>
<td>176</td>
</tr>
<tr>
<td>5.2.3 Data Analysis</td>
<td>179</td>
</tr>
<tr>
<td>5.3 Results</td>
<td>186</td>
</tr>
<tr>
<td>5.3.1 Macrosolute Effects on the Thermodynamics of the Multi-branch Junction</td>
<td>186</td>
</tr>
<tr>
<td>5.3.2 Macrosolute Effects on the Kinetics of RHX Formation</td>
<td>195</td>
</tr>
<tr>
<td>5.4 Discussion</td>
<td>200</td>
</tr>
<tr>
<td>5.4.1 Macrosolute Induced Stability Correlates with Molecularity</td>
<td>200</td>
</tr>
<tr>
<td>5.4.2 Macrosolute Presence Enhances the Rate Constant of RHX Formation</td>
<td>206</td>
</tr>
</tbody>
</table>
5.5 Conclusions ......................................................................................................... 207

5.6 Technological and Physiological Implications of Macromolecular Crowding... 208

5.6.1 Nanostructure Design ................................................................................... 208

5.6.2 Single Nucleotide Polymorphism Detection ................................................ 209

5.6.3 MicroRNA Implications ............................................................................... 211

5.6.4 DNA Replication – Okazaki Fragments....................................................... 213

5.6.5 Resolution of Holliday Junctions ................................................................. 214

References ................................................................................................................. 216
List of Figures

Chapter 1 – Introduction to Nucleic Acids and Macromolecular Crowding

Figure 1.1- The structures of the major purines and pyrimidines ......................... 1
Figure 1.2 - The hydrogen bonding sites in the DNA nitrogen bases .................... 2
Figure 1.3 - The common nucleosides of DNA ...................................................... 2
Figure 1.4 - The syn and anti conformations of adenosine ................................... 3
Figure 1.5 - The adenosine nucleotides ................................................................. 4
Figure 1.6 - Watson-Crick base pairing ................................................................. 6
Figure 1.7 – The least and most destabilizing mismatch base pairs ..................... 15

Chapter 2 - Macrosolute Effects on Bimolecular Systems with Single Mismatches

Figure 2.1- Representative data from the BRCA1-185 thermal denaturation studies 74
Figure 2.2 - Thermal denaturation of the WT duplex in PBS and various macrosolutes ......................................................................................................................................... 75
Figure 2.3 - Thermal denaturation of the WT duplex in the presence of various macrosolutes and an osmolyte .................................................................................................................. 76
Figure 2.4 - Thermal denaturation of the CC duplex in PBS and various macrosolutes ......................................................................................................................................... 77
Figure 2.5 - Thermal denaturation of the CC duplex in the presence of various macrosolutes and an osmolyte ......................................................... 78

Figure 2.6 - Thermal denaturation of the GG duplex in the presence of various cosolutes ................................................................. 79

Figure 2.7 - The relationship between the free energy and the temperature of the duplexes in cosolutes ...................................................... 84

Figure 2.8 - Competition studies in dilute and crowded environments ............. 88

Figure 2.9 - A schematic representation of the distribution of macrosolutes and waters upon duplex formation and dissociation ......................... 94

Chapter 3 - Macrosolute Effects on Bimolecular Systems with Single Base Bulges

Figure 3.1 - Representative wavelength scans and melt profiles of the 5382-WT duplexes in dilute and crowded environments ........................ 119

Figure 3.2 – Representative wavelength scans and melt profiles of the 5382-C duplexes in dilute and crowded environments .......................... 120

Figure 3.3 – Representative wavelength scans and melt profiles of the 5382-T duplexes in dilute and crowded environments .......................... 121

Figure 3.4 – Representative wavelength scans and melt profiles of the 5382-G duplexes in dilute and crowded environments .......................... 122
Chapter 4 - Macrosolute Effects on Unimolecular Systems

Figure 4.1- Representative wavelength scans and the melt profiles of the cyclin D1 stem-loop in PBS and 20% PEG 8K ................................................................. 152

Figure 4.2 - Representative wavelength scans and the melt profiles of the survivin stem-loop in PBS and 20% PEG 8K ................................................................. 153

Figure 4.3 - Representative wavelength scans and the melt profiles for the cyclin D1 duplex in PBS and 20% PEG 8K ................................................................. 156

Figure 4.4 - Representative wavelength scans and the melt profiles for the survivin duplex in PBS and 20% PEG 8K ................................................................. 157

Figure 4.5 - Representative EMSA for the kinetics of hybridization of the cyclin D1 stem-loop to the target-36 strand ................................................................. 159

Figure 4.6 – Identification of the unknown complex band in the cyclin D1 kinetic studies .............................................................................................................. 162
Figure 4.7 - Representative EMSA for the kinetics of hybridization of the survivin stem-loop to the target strand ............................................................................................................. 164

Figure 4.8 – Kinetics of the cyclin D1 and survivin hybridization reactions in the presence of various cosolutes ............................................................................................................. 165

Chapter 5 - Macrosolute Effects on Multi-branch Junctions

Figure 5.1- The sequences and schematic representation of the multi-branch junctions and the RRc duplex ............................................................................................................. 188

Figure 5.2 - Confirmation of the multi-branch structures via EMSA ......................... 189

Figure 5.3 - Representative wavelength scans and melt profiles for the 3-arm structures and the RRc duplex ............................................................................................................. 191

Figure 5.4 - Representative wavelength scans and melt profiles of RHX4B ............... 195

Figure 5.5 - Macrosolute effects on the kinetics of RHX formation .......................... 197

Figure 5.6 - Confirmation of the RHX kinetic product ............................................. 199

Figure 5.7 - A schematic representation of the distribution of macrosolutes and waters upon duplex formation and dissociation ................................................................. 205
List of Tables

Chapter 2 - Macrosolute Effects on Bimolecular Systems with Single Mismatches

Table 2.1 – The extinction coefficients (260 nm) for the BRCA1-185 sequences ..... 61
Table 2.2 – The thermodynamic parameters of hybridization for the WT, CC and GG
BRCA1-185 duplexes in dilute and crowded conditions ............................................ 80
Table 2.3 - The changes in Tm and ΔG_{37C} for the WT, CC and GG duplexes in the
presence of cosolutes ................................................................................................... 81
Table 2.4 - The rate constants of hybridization for the WT, CC and GG duplexes.... 86
Table 2.5 - The change in the preferential hydration (Δν_{21}) upon addition of cosolutes
..................................................................................................................................... 93

Chapter 3 - Macrosolute Effects on Bimolecular Systems with Single Base Bulges

Table 3.1 - The Extinction coefficients (260 nm) for the 5382 sequences .......... 109
Table 3.2 - The thermodynamic parameters of hybridization for the 5382 duplexes in
PBS and macrosolutes ............................................................................................... 124
Table 3.3 - The changes in Tm and ΔG_{37C} for the 5382 duplexes in the presence of
macrosolutes .............................................................................................................. 125
Table 3.4 – The rate constants of formation for the 5382-WT and 5382-C duplexes in
the presence of macrosolutes ..................................................................................... 127
Chapter 4 - Macrosolute Effects on Unimolecular Systems

Table 4.1 - The extinction coefficients (260 nm) for the cyclin D1 and survivin sequences ................................................................................................................... 143

Table 4.2 – The thermodynamic parameters of hybridization for the cyclin D1 and survivin stem-loops and duplexes ................................................................................................................. 154

Table 4.3 - The rate constants of formation for cyclin D1 and survivin duplexes in the presence of various cosolutes ............................................................................................ 165

Chapter 5 - Macrosolute Effects on Multi-branch Junctions

Table 5.1 - Extinction coefficients (260 nm) for the Multi-branch sequences ........ 176

Table 5.2 – The thermodynamic parameters of hybridization for the RR_e duplex, RHX, R*HX and RHX_4B structures ............................................................................................................. 192

Table 5.3 - The change in preferential hydration upon addition of cosolutes ........... 204
List of Schemes

Chapter 1 - Introduction

Scheme 1.1 – The excluded volume effect ................................................................. 22

Chapter 2 - Macrosolute Effects on Bimolecular Systems with Single Mismatches

Scheme 2.1 - The BRCA1-185 duplexes ................................................................. 57

Chapter 3 - Macrosolute Effects on Bimolecular Systems with Single Base Bulges

Scheme 3.1 - The 5382 duplexes .............................................................................. 106

Chapter 4 - Macrosolute Effects on Unimolecular Systems

Scheme 4.1 – The cyclin D1 and survivin sequences ............................................. 141

Scheme 4.2 – The cyclin D1 stem-loop 5’ and 3’ extended sequences ................. 160
Chapter 1 – Introduction to Nucleic Acids and Macromolecular Crowding

1.1 The Chemistry of DNA

Basic Components of DNA

Deoxyribonucleic acid (DNA) is a macromolecule that is made up of nucleotides, specifically 2’-deoxyribonucleotides (1). The three components of a nucleotide are a nitrogen heterocyclic base, a pentose sugar (2-deoxy-D-ribose) and a phosphate group. The sugar and phosphates have structural roles in DNA, whereas the bases are the carriers of genetic information. The nitrogen heterocyclic bases, depicted in Figure 1.1, are grouped into two classes, pyrimidines and bicyclic purines (2). Cytosine (C) and thymine (T) are the major DNA pyrimidines. Adenine (A) and guanine (G) are the major DNA purines.

![Figure 1.1 – The structures of the major purines and pyrimidines.](image)

The bases have atoms and functional groups that can participate in hydrogen bonding (3). In DNA, base pairing takes place through hydrogen bonding between the
bases of the nucleotides. The positions where hydrogen bonds can take place are shown in Figure 1.2.

Figure 1.2 – The hydrogen bonding sites in the DNA nitrogen bases. In duplex DNA, the base pairs are hydrogen bonded together. The hydrogen bond donor sites are shown in blue, and the hydrogen bond acceptor sites are shown in red.

The pentose sugar ring in DNA is 2-deoxy-D-ribose (4). When the bases are attached to the sugar, they are called nucleosides. The names of the nucleosides are deoxyadenosine, deoxyguanosine, deoxycytidine and deoxythymidine for the A, G, C, and T bases respectively (Figure 1.3).

Figure 1.3 – The common nucleosides of DNA.
The purine and pyrimidine bases are connected to deoxyribose through a $\beta$-$N$-glycosidic bond (5). Rotation is possible along the glycosidic bond, which leads to the syn and anti conformations of the nucleosides (Figure 1.4). In the anti conformation, the six membered rings of the purines and the O2 in pyrimidines are pointing away from the sugar. In the syn conformation the purine rings and pyrimidine O2 are pointing over or towards the sugar. Purine nucleosides are in rapid equilibrium of the syn and anti conformations. However, pyrimidine nucleosides are predominantly found in the anti conformation. Generally, the predominant nucleic acid conformation is also anti.

Figure 1.4 – The syn and anti conformations of adenosine. Rotation around the $\beta$-$N$-glycosidic bond (indicated by the arrows) can result in the syn and anti conformations of the nucleotides.

The phosphate esters of the nucleosides are called nucleotides, and they are commonly called the “building blocks” of nucleic acids (6). The common nucleotides contain one to three phosphoryl groups as depicted in Figure 1.5. The nucleotides
contain a deoxyribose and they are called deoxyribonucleotides. In addition to the β-N-glycosidic bond, there are six additional torsion angles which determine the conformation of a nucleotide.

The nucleotides can be linearly linked to form nucleic acids. These nucleotide polymers are joined by phosphodiester bonds (7). The 3’ hydroxyl group of the nucleotide is nucleophilic, and it attacks the α-phosphorous atom of a triphosphate nucleotide, which results in a 3’-5’-phosphodiester and the release of a pyrophosphate. The nucleotide polymer is known as the primary structure of DNA. There are no 5’-5’ or 3’-3’ linkages in common DNA.

Figure 1.5 – The adenosine nucleotides. The adenosine nucleotide is shown containing 5’ phosphate groups: (A) monophosphate (B) diphosphate and (C) triphosphate group.
1.2 The Secondary Structure of DNA

1.2.1 Watson-Crick Base Pairing

In 1953, the first accurate, and currently accepted, model of DNA structure was published by James D. Watson and Francis Crick (8). Watson and Crick were the first to offer information into the base pairing that occurs in DNA. Most DNA in cells are found in a double helical or duplex form which is created by two antiparallel complementary strands of polynucleotides. The individual strands are linked through phosphodiester bonds and the two strands are linked to each other through hydrogen bonding between the nucleotide bases. In Figure 1.2, the hydrogen acceptor and donor properties of the bases are depicted. Watson and Crick discovered that the guanine base paired with the cytosine base through three hydrogen bonds. Adenine and thymine bases were paired through two hydrogen bonds (Figure 1.6). In the 1940’s, Erwin Chargaff and his colleagues made an important discovery, known today as “Chargaff’s rules”, that the ratio of A:T bases was 1:1 and the ratio of C:G was also 1:1 (9). Thus, the discovery made by Watson and Crick was in keeping with Chargaff’s rules.

1.2.2 The Double Helix

Duplex DNA can be pictured as a ladder that is twisted to form a helix. The individual rungs of the ladder are the Watson-Crick base pairs, and the ladder supports are the sugar-phosphate backbones of the DNA (10). The double helix structure results in two unequal grooves called the major and minor grooves (Figure 1.6). The double
helix is stabilized by stacking interactions, hydrophobic effects, hydrogen bonds, and charge-charge interactions (6).

Figure 1.6 - Watson-Crick base pairing. When the base pairs are involved in double helices, a major and a minor groove is created.

Stacking interactions are a result of cooperative, but non-covalent interactions between the upper and lower surfaces of each base that draw the bases together (10). The interior of the helix, where base pairing occurs, is hydrophobic. The exterior, the sugar-phosphate backbone, is hydrophilic. The stacking interactions and the hydrophobic interior result in the twisting of the backbone. The twisting results in the exposure of the sugar-phosphate backbone and the shielding of the hydrophobic bases from water. The hydrogen bonds that form between the base pairs also yield stability to the double helix. Finally, the charge-charge interactions result from the negative charge of the phosphate groups on the backbone and the cations that can be present around a duplex, such as magnesium or sodium. Without the counterbalance of the
cations, the negative charge of the phosphate groups would result in an unstable double helix. There are various conformations of double helix DNA. Three common and well studied conformations of the helix are the A, B and Z form of DNA.

**B-form DNA**

The dominant form of DNA is B-form, a right handed double helix in which the base pairs are stacked above one another and are perpendicular to the long axis of the structure (3). The glycosyl bond conformation is in the anti form and the pentose sugar pucker conformation is C2’-endo. The diameter of the duplex is approximately 20 Å. There are 10.5 base pairs per turn of the helix resulting in a height of 36 Å per full helix turn with a rise per base pair of 3.4 Å. The degree of tilt of the helix relative to the normal of the helix axis is 6°. The B-form is characterized spectroscopically with an ultraviolet absorption maximum at 260 nm and a circular dichroism peak maximum and minimum at 280 and 245 nm respectively (10,11).

**A-form DNA**

A-form DNA is one of the B-form structural variants that is well defined (7). Like the B-form, the A-form is also right handed with the glycosyl bond in the anti conformation. However, in contrast to B-form, the pentose sugar is in the C-3’ endo conformation. The diameter of the A-form is larger, 26 Å, and there are 11 base pairs per turn at a rise of 2.6 Å per base pair. The A-form helix has a 20° base tilt normal to
the helix axis. All of these structural changes in the A-form results in a deeper major groove and a shallower minor groove compared to B-form DNA. If comparing a B-form DNA and an A-form of equal number of base pairs, the A form DNA would appear shorter in height and greater in width. It has been shown that conversion between A and B forms is possible (6). The A-form is also characterized spectroscopically with an ultraviolet absorption maximum at 260 nm. However the circular dichroism peak minimum occurs at 245 nm with no significant peak maximum (6,10).

Z-form DNA

The Z-form is a left-handed helix with a diameter of 18 Å (1). The glycosyl bond and the sugar pucker conformation vary according to the bases. The pyrimidines have an anti glycosyl bond conformation and the sugar pucker conformation is C-2’endo. The purines have a syn conformation for the glycosyl bond and the sugar pucker conformation is C-3’ endo. There are 12 bases per helix rise with 3.7 Å rise per base pair. The bases are tilted 7° normal to the helix axis. As a result of these structural properties, the sugar phosphate backbone of Z-form DNA appears to have a zig-zag conformation. The major groove is almost non-existent; however, the minor groove is narrow and deep. If comparing a B-form DNA and a Z-form, of equal number of base pairs, the Z form DNA would appear longer in height and smaller in width. In cells, there are Z-DNA tracts, commonly made up of alternating G and C residues (6).
1.2.3 Other Structural Forms of DNA

In addition to the double helix, DNA can exist in a variety of structural forms. Palindromes are common in DNA, a sequence which can be read the same backwards or forwards (7). For example, a single strand of DNA can form a stem-loop structure as a result of palindrome regions. The stem typically ranges from 3-6 base pairs on average, and the loop size can vary from 4 to 20 or more bases (12). A cruciform is another palindrome based structure (13). Cruciform structures can be found when DNA underwinds.

A triple helix is possible when a third strand of DNA binds to a Watson-Crick duplex (14). The base pairing between the second and third strand is called Hoogsteen pairing (7). A protonated cytosine can bind to the guanine residue of a G-C base pair; likewise, a thymine can pair with the adenine of an A-T base pair. Tetraplex DNA are also called quadruplexes, and they can form in G-rich DNA (15). C-rich DNA can form the I-motif structure, which is created by two antiparallel-stranded duplexes bound in a parallel manner to each other (16). In this structure, the base pairing occurs between a cytosine and protonated cytosine base. DNA is also capable of forming multi-branched junctions, where three or more complementary strands comprise the final structure. These multi-helix junctions are sometimes referred to as multi-arm junctions. Involving only canonical base pairing, 3, 4, and 5-arm junctions have been reported (17,18). Four-way junctions, or Holliday junctions, in particular, are well studied due to their importance in homologous recombination (19).
1.3 Hydration of DNA

The amount of water associated, or the extent of hydration of DNA, is important to the stabilization of the double helix structure (20). The information available on the hydration of DNA stems mostly from the Dickerson-Drew dodecamer experiments. Dickerson and colleagues studied over 68 variations of the sequence with X-ray crystallography (21). The B-DNA structure is the most hydrated of the three forms previously described (see section 1.2.2). In B-form DNA, there are approximately 14-17 highly ordered waters per base pair. The major groove contains two hydration layers. The first layer contains approximately 19 water molecules which solvate the backbone and interact directly with exposed carbonyl groups, nitrogen and oxygens. The second hydration layer spans the entire major groove (10). The minor groove is where the “spine of hydration” is formed. A zigzag pattern of water characterizes the spine of hydration. Alternating water molecules are in direct contact with the bases, particularly with A-T base pairs. Upon dissociation of a B-DNA helix, approximately four water molecules are released (22).
1.4 Single Base Mutations in DNA

1.4.1 Mismatched Bases

There are two types of base pair mismatches: transition and transversion mismatches (10). Transition mismatches are those in which a purine is paired with the incorrect pyrimidine. Transversion mismatches occur when two purines, or two pyrimidines, are paired together. In addition to the type of mismatch, it is important to understand how the double helical structure is affected by the mismatch. However, single base mismatches can be difficult to detect. The difference in free energy between a wild-type and single base mismatch mutant, is about 3-6 kcal mol\(^{-1}\) (23).

**Transition Mismatches**

The two possible transition mismatches are a GT and AC base pair (10). Both of the mismatches are wobble pairs, and contain two hydrogen bonds between them. Crystal structure studies of oligonucleotides have provided evidence of the GT base pair transition mismatch in A, B, and Z conformations. In the GT base pair, the glycosidic bonds are in the *anti-anti* conformation. The backbone of the oligonucleotides containing GT mismatches were not distorted at all when compared to the wild-type oligonucleotide structure.

In the AC base pair, protonation of the N1 in adenine results in base pairing through hydrogen bonds (6). Interestingly, it has been found that no water molecules are associated with the AC base pairs in the minor groove. In contrast, in wild-type
duplexes, the AT base pair is heavily hydrated within the minor groove. Comparison of wild-type to AC mismatched duplexes has shown that the base stacking and helix conformation are only slightly perturbed.

**Transversion Mismatches**

There are six possible transversion mismatches, AA, GG, CC, TT, GA, and CT (10). The GA mismatch can adopt two conformations $G^{(anti)}$-$A^{(anti)}$ and $G^{(anti)}$-$A^{(syn)}$, and both contain two hydrogen bonds (24). The $G^{(anti)}$-$A^{(syn)}$ causes little disruption to the structure of B-DNA. However, the $G^{(anti)}$-$A^{(anti)}$ is more disruptive as a result of the increased distance between the bases. $G^{(anti)}$-$A^{(anti)}$ base pairing results in changes of hydration and broadens the minor groove of B-DNA (6).

The homo transversion mismatches vary in their base stacking potential. The GG mismatch has two hydrogen bonds, and their purine rings are positioned in a way that high stacking potential is possible (25). The GG mismatch has been found to have a *syn-anti* conformation as confirmed by crystal structure determinations (26). However, the GG mismatch was also found to be in rapid equilibrium between *syn-anti* and *anti-syn* (27). The introduction of a GG base pair into a duplex has also been shown to increase the helical width and narrow the minor groove. Another mismatch with high stacking potential is the AA base pair, but only one hydrogen bond stabilizes the structure. Low stacking potential is a characteristic of both CC and TT mismatches.
However, TT mismatches are stabilized by two hydrogen bonds, whereas CC is stabilized by only one hydrogen bond (25).

**The Stability of Transversion and Transition Mismatches**

The stability of a DNA duplex can be measured via the melting temperature, $T_m$, which is dependent on four factors: strand length, strand concentration, base sequence, and ionic strength of added salt (28). In general, strand length, strand concentration and the ionic strength are directly proportional to duplex stability. Of the Watson-Crick base pairs, GC and AT have been thermodynamically studied, and it has been reported that DNA with GC or CG base pairs were more stable than those with AT or TA pairs in all cases examined (6). This is due to the hydrogen bonding between the two base pairing schemes; GC contains three hydrogen bonds and AT only contains two.

The stability of transition and transversion mismatches has also been studied in detail. It has been shown that the general order of stability within a duplex with transversion mismatches, specifically the homo-mismatches, is GG $>$ AA $\sim$ TT $>$ CC as reported by Chalikian and colleagues (25). Identical results were observed by Santalucia and colleagues (24). The stability scale for both transition and transversion mismatches was reported by Gaffney and Jones who studied the thermal stability of various duplexes (29). Gaffney and Jones concluded that the order of least to most destabilizing mismatches on duplexes is as follows: GT $>$ GG $=$ AG $>$ TG $>$ GA $=$ TT $>$ TC $>$ AC $>$ CT $>$ AA $>$ CA $>$ CC. Similar, but not identical, ranking order of stability
have been reported by Ke and Wartell, who studied the mismatches in various sequence contexts (30). Although Ke and Wartell had four stability rankings, all of their stability schemes were in general agreement with each other as well as with previous findings by Gaffney and Jones. Finally, Santalucia and colleagues also studied the effects of mismatches on duplex stability (24). They reported the following general trend: GC > AT > GG > GT ~ GA > AC⁺ > TT ~ AA ~ CC⁺ > TC ≥ AC ≥ CC.

Overall, it can be seen that GA and GG mismatches are consistently one of the most stable, and thus, least destabilizing to the duplex. Conversely, CC is the least stable and hence the most destabilizing to the duplex structure. GG mismatches are the least destabilizing to a duplex structure because of the high stacking potential of their purine rings (Figure 1.7). Also, GG base pairs have two hydrogen bonds. CC base pairs are the most destabilizing to a duplex because of their low stacking potential and their hydrogen bonding to each other consists of only one hydrogen bond (Figure 1.7). Additionally, the nature of the mismatch stability is also dependent upon nearest neighbor interactions.
Figure 1.7 – The least and most destabilizing mismatch base pairs. Of all the possible mismatches, the CC and GG transversion mismatches are the most and least destabilizing to duplex structure. (A) GG base pair and (B) the CC base pair.

Interestingly, in vivo studies of post-replication repair of mismatched bases have shown that GG and GT mismatches are the most efficiently repaired (31-34). GA and CC mismatches are the least efficiently repaired. The repair efficiency does not appear to be related to the thermal stability of the mismatches since GA is more stable than CC, yet they both have poor repair rates.
1.4.2 Bulge Bases

In addition to single base mismatch mutations, DNA can also have insertion or deletion of a single base. As a result of single base deletion/insertions, one of two things can occur; the incorporation or exclusion of the base (6). Single base insertions/deletions are also called frame shift mutations because the removal or addition of a base can shift the reading frame. If the mutation occurs, for example, in double helix DNA, and the helix cannot accommodate the mutation, a “bulge” will form and the base is said to be extrahelical. The bases around the extrahelical bulge can still be involved in complementary base pairing, and the structure may not be significantly affected. A second scenario can take place when the single base that is deleted or inserted can be accommodated into the structure of the DNA, and thus is said to be intrahelical. If the secondary structure is a double helix, the bases can shift by one and still accommodate the intrahelical base without major disruption to the helix. However, it is possible that by shifting by one base, transition or transversion mismatches may occur.

Unlike the mismatched base pair, there is no stability scale for the bulge bases because their extent of helix disruption is dependent on the identity of both the bulge base and neighboring bases (35). However, some general trends have been noted for bulge bases. For example, Wartell and Ke categorized the bulge bases into two groups based on their work on single bulge base effects on duplexes (36). Group I classification is given to bulges that are of different identity than its neighboring bases.
Group II classification describes bulges that are identical to at least one of its neighboring bases. Wartell and Ke found that group II bases are not as disruptive to the helix as bulge bases in group I. Furthermore, Wartell and Ke noted that in general pyrimidines bulges are more disruptive than purines in helix. The work of Woodston and Crothers reaffirms the findings put forth by Wartell and Ke (37). The differences in stabilization are attributed to positional degeneracy of group II bulges (36). Additionally, it is believed that aside from the group I versus group II classification, the conformational freedom of the purine rings versus pyrimidine rings is also a factor in how they affect duplex stability (38).

1.5 The Crowded Intracellular Environment

The intracellular environment contains a variety of macromolecules that collectively occupy 20-30% of the cell’s available volume which creates a crowded environment (39). Crowding is a universal property of all types of cells and can be found in numerous biological systems. Various macromolecules such as proteins, nucleic acids, and polysaccharides crowd the cytoplasm of prokaryotic cells (40). In *Escherichia coli*, the average concentration of total RNA in a cell is 75-150 g L⁻¹ and the concentration of proteins is 200-300 g L⁻¹. The resulting combined concentration of macromolecules is thus 300-400 g L⁻¹ (41). One of the macromolecules that crowds red blood cells is hemoglobin, present at an average 350 g L⁻¹ (42). However, crowding
is not restricted to the interior of a cell; crowding is also present outside the cell. For example, blood plasma is crowded with 80 g L\(^{-1}\) of protein (43).

Biological systems have evolved to function in crowded environments; however, biological experiments are typically conducted in dilute systems (44). The more common properties that are regulated in biological experiments are pH, temperature and salt concentrations (45). However, crowding effects can be substantial and should not be ignored, but rather taken into account along with other conditions (46). A crowded environment can result in significant changes in thermodynamic and equilibrium values. Crowding effects arise as a result of changes in the available volume and through changes in the distribution of the water and crowding molecules. The excluded volume effect is used to describe the changes in available volume. Here, the Kirkwood-Buff theory of solutions is used to interpret changes in water and crowding molecule distribution.

1.6 Mimicking a Crowded Environment

Synthetic Substitutes for Intracellular Macromolecules

In order to mimic the crowded intracellular environment in a reproducible and controllable manner, synthetic, inert substitutes are used to replace the cellular macromolecules (47). The substitutes are collectively termed “cosolutes” and are classified according to their molecular weight. Macrosolutes are cosolutes whose
molecular weight is greater than 1000, such as Ficoll Type 70 (70 kDa) or large poly(ethylene glycol) (PEG). Osmolytes refers to cosolutes whose molecular weight is less than 1000, such as glycerol, ethylene glycol, and saccharides.

Size and Shape Dependence of Crowding Effects

In systems dominated by size effects versus hydration, the larger the volume excluded to a molecule of interest, the higher the thermodynamic activity may increase, and a greater crowding effect may be observed (39). For example, osmolytes such as glycerol tend to have negligible or negative effects on excluded volume as compared to larger cosolutes like PEG 8000 (22). Thus, macrosolutes are preferred to osmolytes in order to attain the greatest possible crowding mediated effect.

Size and shape are also important factors to consider when mimicking crowding in the laboratory. Globular or circular shapes have been shown to exert an increased effect on a nucleic acid system as compared to rod-like structures (48). The size of the cosolute relative to the test species is also crucial (49). The excluded volume effect is greatest when the unoccupied volume between cosolutes is small and inaccessible to the molecule of interest, thereby reducing the available volume. If a molecule of interest is the same size or larger than the cosolute, there may be crowding effects (50). Therefore, both the relative and absolute size of the cosolute are important for attaining excluded volume effects.
Additional Cosolute Selection Criteria

The selection of cosolutes does not only depend on the factors described above; various additional criteria should be considered (39). First, it is important that the cosolute does not interact with the system, with the exception of steric repulsion. Second, the cosolutes should be water soluble and not susceptible to self-aggregation. To prevent solutions from being too viscous, an ideal crowding agent should be globular and not extended. Also, it is important to use cosolutes with a wide molecular weight range in order to separate geometric effects from osmotic effects. In a study by Zimmerberg and Parsegian, PEGs of various sizes affected conductance through alamethicin channels by lowering the water activity (51). However, smaller PEGs were found to affect the protein channel by physical blockage of the channel as a result of the PEGs being small enough to penetrate the channels.

Validating Cosolute Effects

In order to determine if the effects observed in a crowded system are due to crowding or to specific interactions between particles, certain criteria must be confirmed (50). First, the effects observed with one cosolute should be duplicated with different cosolutes. Second, if the concentration of the cosolute is increased, a corresponding increase in effect should be observed. Third, larger cosolutes should cause a larger effect than smaller cosolutes, when compared at the same number density. However, the opposite is expected when they are compared at the same weight
concentration. In addition, it should be noted that all other factors of a system are kept constant when investigating crowding effects, such as pH, salt concentrations, and others.

1.7 Crowding Effects - Excluded Volume Theory

The term “concentrated” is not used to describe the cell’s interior because no single macromolecule is present at a high concentration (52). Crowding refers to the collective fractional occupancy of the macromolecules. As a result of crowding, the accessible volume in a cell becomes reduced or excluded. The exclusion characteristic is why crowding effects result in part from the “excluded volume effect.” Crowding is a physical non-specific effect that arises in part from the steric repulsion in highly volume-occupied systems (43). The nonspecific interactions between the macromolecules depend on global properties of the macromolecules such as net charge, polarity of the surface residues, and shape (52). Nonspecific interactions can be repulsive (steric and electrostatic) or attractive (hydrophobic and electrostatic) and are generally weaker than specific interactions. Macromolecular crowding is a type of steric interaction (53). The nonspecific interactions between the macromolecules contribute to the free energy of the biological system which has significant consequences on rates, thermodynamics, and equilibria (54,55).
Scheme 1.1 – The excluded volume effect. (A) A test species, 1, that is smaller than the background species is introduced into the system. Here, the activity coefficient is approximately 1 and the thermodynamic activity is approximately equal to the concentration. Conversely in (B) a test species, 2, that is equal to the background molecules is introduced into the system. Here, the activity coefficient is greater than one, resulting in a chemical activity greater than the concentration. This scheme is modeled after figure 2 from Allen P. Minton (52)

Excluded volume effects can be demonstrated by Scheme 1.1, where macromolecules, represented by the blue spheres, occupy 20% of the available volume in a “cell.” The volume that the macromolecules take up within this cell is inaccessible to any other species. In Scheme 1.1, the molecule being introduced into the system is called a test molecule, and the macromolecules are called background molecules. In part A, test species 1 is introduced into the system, and is much smaller than the background molecules (56). Therefore, test species 1 will be able to access virtually all of the available volume that is not being occupied by the background molecule, all of the area shown in gray. In part B however, the test species is larger than the background molecules. As a result, the volume available to species 2 will be
significantly reduced as compared to the volume available to species 1. The available volume to species 2 is represented by the area outside of the dashed lines. The reason for the difference in available volume is because the center of a molecule being introduced can approach the center of the other molecules to no less than the distance at which the surfaces of the two molecules meet. This restricted distance is indicated by the dashed lines around the background species. Hence, the volume available in a system will depend on the size of the molecule that is trying to occupy the space.

The background molecules and test species in Scheme 1.1 interact with each other only non-specifically through steric interactions. In a system such as this, the activity coefficient, $\gamma_i$, is equal to the effective concentration, $a_i$, over the actual concentration, $c_i$, as shown in equation 1.1.

$$
\gamma_i = \frac{a_i}{c_i} = \frac{v_{total}}{v_{avail}}.
$$

(Equation 1.1)

$$
a_i = c_i
$$

(Equation 1.2)

Under ideal conditions, when the finite size of a molecule has no effect, equation 1.1 can be reduced to equation 1.2 where the $\gamma_i$ is approximately equal to 1, this is most commonly assumed and allows for the use of concentration for convenience. In a crowded environment, which can be described as non-ideal, solute-solute interactions cannot be neglected and $\gamma_i$ cannot be reduced to 1 because the chemical activity of a species in a crowded solution is greater than 1. As the concentration of the cosolute increases, the activity coefficient increases non-linearly, resulting in increased crowding.
effects. Although crowding can increase the activity of a species, it also decreases the rate of diffusion (57). While reaction rates can be increased by crowding, they must decrease at some point as they become diffusion limited.

In a biochemical association reaction, crowding favors the state that excludes the least volume: the more compact state (49). In a dissociation reaction, such as duplex DNA melting, the associated state is favored by crowding. For a change of state reaction, such as a protein unfolding, crowding will tend to favor the folded, more compact state, over the unfolded state.

1.8 Crowding Effects - The Distribution of Waters and Cosolutes

Water is at the center of many biomolecular processes such as physiological regulation, protein folding, protein-ligand interaction, protein denaturation, DNA-ligand interactions, and DNA hybridization (58-61). Biomolecular processes, such as the association and dissociation of macromolecules, change the hydration of the system through the release or absorption of water molecules (59,62). For example, hemoglobin absorbs 60 water molecules as it transitions from its deoxygenated to oxygenated form (63). Hexokinase closes down on glucose in an interaction that results in the release of 320 water molecules (64).

The exact role that water plays in these processes can vary. Water molecules can act as bridges between macromolecules in a specific interaction with the surface of
the macromolecule, or they can also be found trapped inside cavities within the macromolecule (65,66). The last 10-15 Å of separation between surfaces contains water molecules whose structure is believed to dominate the interaction between the two macromolecules (67,68). Therefore, to gain a better understanding of the mechanisms in biological processes, the water molecules and their structure must be taken into account (69).

The stability of macromolecules can be altered through cosolute addition, as will be described in detail in section 1.9. For example, DNA can be stabilized by the addition of macrosolutes or destabilized by the addition of osmolytes (70,71). Proteins however, can also be stabilized by both osmolytes and macrosolutes, or denatured by urea and guanidine hydrochloride (72,73). The variety of cosolutes effects produced by crowding warrants a universal theory to interpret how cosolute effects arise. Such a theory must be able to interpret effects for cosolutes of all sizes. Osmotic stress and the Kirkwood-Buff theory have been at the center of the debate regarding how cosolutes mediate their effects (67,69,74). Proponents of osmotic stress theory claim that crowding effects (from macrosolutes and osmolytes) are based on changes in hydration (67,75,76). On the other hand, proponents of Kirkwood-Buff theory claim that macrosolute effects are cosolute exclusion mediated while osmolyte effects are mediated by changes in hydration (69,74,77).

The following sections describe these theories for a ternary system where 1=water, 2=macromolecule, and 3=cosolute.
1.8.1 The Local Bulk Partitioning Model (LBPM)

In order to understand the interactions between water, cosolutes and macromolecules, the location and distribution of these molecules in solution and on the macromolecule’s surface is important. According to the local bulk partitioning model (LBPM), there are two general locations where water and cosolutes can reside; the bulk or local domain (78,79). The “bulk domain” is a region of solution that is sufficiently distanced from the solute such that in this region, the solvent has characteristics of the bulk solvent. The “local domain” is where the solvent properties have been altered because of the solute. There is no clear definition of where the “local domain” ends around a macromolecule, and where a “bulk domain” begins. However, it is generally accepted that at least the first two solvation shells are part of the “local domain.”

1.8.2 The Preferential Hydration Parameter

Osmotic stress and Kirkwood-Buff theories both employ the preferential hydration parameter. The preferential hydration parameter, $\nu_{21}$, describes the effect that the water (in terms of chemical potential), $\mu_1$, has on a macromolecule, $\mu_2$, at a constant pressure, $P$, temperature, $T$ and molality of the macromolecule, $m_2$ (80).

$$\nu_{21} = \left(\frac{\partial \mu_2}{\partial \mu_1}\right)_{P,T,m_2}$$  (Equation 1.3)

Equation 1.3 can be re-written to describe conditions at a constant chemical potential of water, $\mu_1$. 

26
\[ v_{21} = \left( \frac{\partial m_2}{\partial m_1} \right)_{P,T,\mu_1} \]  

(Equation 1.4)

This equation describes the number of water molecules associated with the addition of a macromolecule (81). The preferential hydration parameter can also be expressed in terms of the number of water, \( N_{21} \) and cosolute, \( N_{23} \), that are interacting with the surface of the macromolecule (82). Specifically, \( N_{2i} \) is the number of component \( i \) in the hydration shell. Hence, the structure of the solution around the macromolecule is described by the preferential hydration parameter.

\[ v_{21} = N_{21} - \frac{m_1}{m_3} N_{23} \]  

(Equation 1.5)

When a macromolecular reaction takes place, such as hybridization of DNA, the change in the preferential hydration parameter, \( \Delta v_{21} \), can be expressed as the change in the number of cosolutes and water distributed around the surface of the DNA, \( \Delta N_{23} \) and \( \Delta N_{21} \) respectively (77).

\[ \Delta v_{21} = \Delta N_{21} - \frac{m_1}{m_3} \Delta N_{23} \]  

(Equation 1.6)

### 1.8.3 Osmotic Stress Analysis

Osmotic stress (OS) analysis treats water as a ligand in macromolecular processes. This method focuses on enumerating the number of water molecules absorbed or released as a result of a macromolecular process (76). OS measurements
are performed in the absence and presence of osmolytes (58,64,83). The change in the number of water molecules, and the corresponding change in the chemical potential of water, with and without osmolyte, reflect a change in the water activity (67). Although osmotic stress is a popular theory/analysis, its accuracy, validity, and applicability to the effect of solutes has been the source of recent debate (47,75,78,79,84). Specifically, the use of osmotic stress to interpret macromolecular crowding effects has been the focus of several papers (22,69,74,75,77,80,84-87).

According to OS analysis, the dominant contribution to $\Delta \nu_{21}$ is the change in hydration, $\Delta N_{21}$, when cosolutes are added (63,67,76). However, this statement is accompanied by an assumption that cosolutes do not associate with the macromolecule surface as a result of volume exclusion (63,67,76). This assumption results in $\Delta N_{23}=0$, where the preferential hydration parameter is equal to the number of water molecules associated with the surface of the macromolecule (77). As a result equation 1.6 is rewritten as:

$$\Delta \nu_{21} = \Delta N_{21} . \quad \text{(Equation 1.7)}$$

However, equation 1.7 is only valid when the cosolute is completely excluded from the hydration shell of the macromolecule (78). Consequently, only strongly excluded osmolytes, such as betaine, would result in the validity of equation 1.7 (74,77,78). The assumption of zero association is inaccurate for macrosolutes, and some osmolytes, and results in an overestimation in the changes in hydration, $\Delta N_{21}$ (69). Hence, the
accuracy of OS analysis is highly dependent on the osmolyte selection. It has been
demonstrated that OS measurements do not support the behavior of larger macrosolute
effects (69,74,77,85). Also, OS analysis does not rely on structural information for its
analysis (67,69). Due to these limitations and disadvantages, OS analysis is only valid
for systems containing certain osmolytes, and not macrosolutes.

1.8.4 The Kirkwood-Buff Theory of Solutions

In 1951, John G. Kirkwood and Frank P. Buff published a paper describing a
general statistical mechanical theory of solutions (88). The Kirkwood-Buff (KB) theory
demonstrated a relationship between the integrals of radial distribution function of
molecular pairs in the system to the derivatives of chemical potentials and osmotic
stress (based on concentrations), partial molar volumes, and compressibility. Over the
last 20 years newer, matrix free, derivations of KB theory have led to numerous
applications (85,89-94).

Overview of Kirkwood-Buff Theory

The preferential hydration parameter, $\nu_{21}$, represents the excess water versus the
cosolute that is interacting with the surface of the macromolecule (81). Conversely, the
preferential cosolute interaction parameter, $\nu_{23}$, represents the excess cosolute versus the
water that is interacting with the surface of the macromolecule (80). In KB theory $v_{21}$ and $v_{23}$ are defined as:

$$v_{21}^\sigma = -\left(\frac{\partial \mu_2^\sigma}{\partial \mu_1}\right)_{T,P,m_2} = N_{21} - \frac{m_1}{m_3} N_{23} \quad (Equation \ 1.8)$$

$$v_{23}^\sigma = -\left(\frac{\partial \mu_2^\sigma}{\partial \mu_3}\right)_{T,P,m_2} = N_{23} - \frac{m_3}{m_1} N_{21} \quad (Equation \ 1.9)$$

where $m_i$ represents the molarity (density) of species $i$, and $N_{21}$ and $N_{23}$ represent the excess solvation number of water and cosolute around the macromolecule, respectively (74). In equation 1.8 and 1.9, both parameters are defined for a particular state of the macromolecule (represented by $\sigma$). The relationship between $v_{21}$ and $v_{23}$ is described by using the molarity of water and the cosolute (80).

$$v_{21}^\sigma = -\frac{m_1}{m_3} v_{23}^\sigma \quad (Equation \ 1.10)$$

The excess solvation numbers, $N_{21}$ and $N_{23}$ are related to the partial molar volume of the macromolecule, $V_2$, where the isothermal compressibility of the solvent mixture, $\kappa_T$ and $k$, the Boltzmann distribution constant, are negligible for proteins (85).

$$V_2 = -V_1 N_{21} - V_3 N_{23} + kT \kappa_T \quad (Equation \ 1.11)$$

$$N_{2i} = m_i N_A \int d\mathbf{r} [g_{2i}(\mathbf{r}) - 1] \quad (Equation \ 1.12)$$
$N_{2i}$ is defined by $g_2(r)$ which is a radial distribution function that describes the density of component $i$ at a particular distance, $r$, from the macromolecule, $N_A$ is Avogadro’s number and $m_i$ is the mass of particle $i$ (85). Excess solvation numbers, $N_{2i}$, have contributions from solvent reorganization and the excluded volume that arises because the solute cannot access the core of the macromolecule. The number of bound solvent molecules, $N'_{2i}$, is given by:

$$N'_{2i} = N_{2i} + m_i V_E .$$  \hspace{1cm} \text{(Equation 1.13)}

In order to estimate the excluded volume of a macromolecule, $V_E$, structural information is needed. However, equation 1.13 is most accurate for systems where the macromolecule is globular because $V_E$ is estimated using a hard-sphere model approach. Hence, this model may not be as accurate for macromolecules such as DNA. Although OS measurements and LBPM attempt to define the solvation shell boundaries, Kirkwood-Buff contains no assumptions of where the solvation shells are or where their boundaries are. For Kirkwood-Buff analysis, the solvation shell range is not relevant (69).

KB has many specific advantages over OS analysis since it can be applied to cosolutes of any size, shape or complexity (90). KB can also be applied to a system with any number of components. Unlike OS analysis, KB does not make approximations on solvation shell boundaries, and utilizes structural information for its analysis in order to estimate the number of water molecules (95).
Kirkwood-Buff versus Osmotic Stress Analysis

The number of water molecules associated with oxygen uptake by hemoglobin and the binding of camphor to cytochrome P450 have been analyzed extensively by OS (63,96). Recently, both hemoglobin and cytochrome P450 reactions were evaluated by OS and KB analysis where direct comparisons were made regarding the number of water molecules associated with the reactions (69).

In order to calculate the number of waters involved in the hemoglobin reaction, all possible conformational changes when hemoglobin takes up oxygen were used to obtain $\Delta V_E$, the change in volume exclusion. OS analysis determined that approximately 65 waters were absorbed for all five possible conformations. KB analysis results had a range of -65 to +20 water molecules absorbed for all five possible conformations. OS analysis showed that 19 waters were absorbed in the cytochrome P450 camphor binding reaction, whereas KB analysis determined that only 7 waters were absorbed.

According to the KB analysis, which differentiated between bound and associated waters and did not assume that the $\frac{m_1}{m_3} N_{23} = 0$, the hemoglobin and cytochrome P450 systems OS analysis overestimated the number of water molecules absorbed in a reaction. The OS overestimation stems from the restriction of the hydration parameter (equation 1.5). The zero association assumption requires the
\( \frac{m_1}{m_3} N_{23} \) term to be equal to zero, however, in the examples shown here, the value of \( \frac{m_1}{m_3} N_{23} \) was a large negative number.

### 1.8.5 Applications of Kirkwood-Buff Theory to Crowded Systems

#### Determination of \( v_{21} \) in Crowded Systems

KB theory was applied to investigate the effects of cosolutes on macromolecules (77). The hydration parameter, \( v_{21} \) was calculated in order to determine whether the dominant contributor was hydration, \( N_{21} \), or cosolute effects, \( N_{23} \). Two proteins, ribonuclease A and bovine serum albumin (BSA) were placed in crowded solutions. In the presence of various macrosolutes (PEGs of average molecular weight of 2000 to 6000), the \( v_{21} \) was calculated for each of the proteins. The calculations showed that the \( v_{21} \) for ribonuclease A and BSA was dominated by \( \frac{m_1}{m_3} N_{23} \), a large negative number, in the presence of all the macrosolutes.

A negative \( \frac{m_1}{m_3} N_{23} \) indicates that the PEG molecules are being excluded from the surfaces of the proteins (see equation 1.6). Although OS analysis assumes that \( v_{21} \) is dominated by hydration and not cosolute exclusion, the opposite effect is observed for ribonuclease A and BSA in PEGs. As the PEG MW increased, the hydration contribution to \( v_{21} \) decreased. In addition, this analysis on BSA and ribonuclease
demonstrated that there is no correlation between hydration and the exclusion of cosolute.

The stability of ribonuclease A was investigated in the presence of trehalose, sucrose, glucose, and glycerol to explore how osmolytes effects can be interpreted (77). The results, which were analyzed by KB theory, indicated that the preferential hydration parameter was not dominated solely by the exclusion of the osmolytes. Additionally, the contribution from hydration was not negligible. The analysis on ribonuclease A with macrosolutes and osmolytes led to the conclusion that macrosolute effects are dominated by cosolute exclusion and osmolyte effects are dependent on both cosolute exclusion and hydration changes.

**Kirkwood-Buff Theory Analysis of Δν21 in Crowded Systems**

A reaction, such as $A \rightarrow B$, in a crowded system can be described by $\Delta \nu_{21}$, the change in the preferential hydration parameter and $\Delta \nu_{23}$, the change in the preferential cosolute interaction parameter.
\[ \Delta \nu_{21} = RT \left( \frac{\partial \ln K}{\partial \mu_1} \right)_{T,P,n_2} \]  \hspace{1cm} (Equation 1.14)

\[ \Delta \nu_{23} = RT \left( \frac{\partial \ln K}{\partial \mu_3} \right)_{T,P,n_2} = \Delta N_{23} - \frac{m_3}{m_1} \Delta N_{21} \]  \hspace{1cm} (Equation 1.15)

\[ \Delta N_{21} = n_1 N_A \Delta G_{2i} = n_1 N_A \left( G_{2i}^B - G_{2i}^A \right) \]  \hspace{1cm} (Equation 1.16)

\[ G_{2i}^B = \int d\vec{r} \left[ g_{2i}^B(\vec{r}) - 1 \right] \]  \hspace{1cm} (Equation 1.17)

The KB parameter, \( G_{2i}^B \), is defined by the correlation function \( g_{2i}^B(\vec{r}) \). The integral expression in equation 1.17 represents the difference in the distribution of species \( i \) around the macromolecule versus a bulk domain solution (92). The KB parameter \( G_{2i} \) describes the macrosolute-water solvation, and \( G_{23} \) describes the macrosolute-cosolute solvation. Thus, \( \Delta G_{2i} \) defines the preferential interaction of the macromolecule and species \( i \) (97).

In order to determine whether the preferential hydration parameter is dominated by hydration, \( G_{21} \), or solute exclusion, \( G_{23} \), researchers have used the data available for several macromolecular systems and analyzed the effects of cosolutes through KB theory (74). Shimizu and Boon analyzed several systems including; (1) hexokinase-glucose dissociation in the presence of PEGs (MW up to 10000) and (2) lysozyme denaturation by guanidine hydrochloride (GuHCl). For each of these systems, \( \Delta \nu_{21} \), \( n_1 \Delta V_2 \), \( \Delta G_{21} \), \( \Delta G_{23} \) and \( \Delta V_E \) were tabulated.
The cosolute induced modulation of glucose-hexokinase dissociation showed that the change in hydration parameter, $\Delta \nu_{21}$ (326 mol mol$^{-1}$) was dominated by $\Delta G_{23}$ (-17812 ml mol$^{-1}$) because the contribution from $\Delta G_{21}$ was smaller (232.4 ml mol$^{-1}$). The large negative value of $\Delta G_{23}$ indicated that the PEG molecules were being excluded from the surface of the dissociated protein. As expected, the dissociated form of the protein had a larger $V_E$, as indicated by the $\Delta V_E$ (+521 ml mol$^{-1}$). The exclusion of PEG from the dissociated form of hexokinase results in a stabilization of the associated hexokinase-glucose. The magnitude and sign of $\Delta G_{23}$ and $\Delta G_{21}$ demonstrate that crowding effects are not dominated by hydration changes, as OS theory indicated, but rather the exclusion of cosolutes was the major contributor to the observed effect.

The second system studied by Shimizu and Boon, was the denaturation of lysozyme by GuHCl. In contrast to the hexokinase system, where a stabilization effect was observed, here the system was destabilized and denaturation was observed. The lysozyme denaturation yielded a $\Delta G_{23}$ (+2123 ml mol$^{-1}$) and a $\Delta G_{21}$ (+55 ml mol$^{-1}$). In this case, the contribution from the hydration effects cannot be dismissed. As a result, the effects noted in the lysozyme denaturation are mediated through both hydration changes and cosolute exclusion. The positive $\Delta G_{23}$ described the GuHCl as being associated with the protein in the denatured state, and hence excluded from the native state. Therefore the lysozyme is driven towards the denatured state. A positive value was obtained for $\Delta V_E$ because the protein is being denatured, and so its unfolding leads to an increase in the excluded volume. These results support the view of denaturants as
being preferentially accumulated on the surface of the associated macromolecule and thus driving the denaturation reaction (97). The destabilization of systems by osmolytes can be theorized to function in the same manner as denaturants. Shimizu and Boon also proved that when the absolute value of the change in the hydration parameter was much greater than the change in volume of the macromolecule times the molarity of water \(|\Delta V_2| \gg |n_1 \Delta V_1|\), the change in hydration parameter \((\Delta \nu_{21})\) will be dominated by the macrosolute changes in distribution \((\Delta G_{23})\). Thus, the preferential parameters will be dominated by the distribution of the cosolutes, and not by water.

**Kirkwood-Buff Theory and Crowding Conclusions**

The results obtained by Shimizu and Boon’s study of hexokinase and lysozyme shed light on how macromolecular crowding effects are mediated. Based on KB analysis, crowding effects by macrosolutes are mediated through solute exclusion \((\Delta G_{23})\) and effects of osmolytes and denaturants are mediated through both solute exclusion and hydration changes \((\Delta G_{21})\). With hexokinase, a crowding stabilized system, the large negative value for \(\Delta G_{23} (-17812 \text{ ml mol}^{-1})\) and the small positive value for \(\Delta G_{21} (+232.4 \text{ ml mol}^{-1})\) indicated that hydration contributions were negligible. Hence, for hexokinase the major contribution to the system was the cosolute exclusion. Conversely, for lysozyme the \(\Delta G_{23} (+2123 \text{ ml mol}^{-1})\) value was positive and small, but the \(\Delta G_{21} (+55 \text{ ml mol}^{-1})\) could not be disregarded. Thus, for a system with an osmolyte or denaturant, both the hydration and solute exclusion contributions are important.
The KB analysis performed by Shimizu and Boon are in agreement with the findings of Abraham Minsky on DNA duplex dissociation in the presence of cosolutes (71). Minsky noted that a 4.5% ethylene glycol solution and a 15% PEG 8K solution both modulate the water activity (a change in hydration) to the same extent. However, the duplexes were stabilized against thermal denaturation in 15% PEG 8K, destabilized in 15% ethylene glycol and no effect was noted for the 4.5% ethylene glycol. Hence, the size of the cosolute determined whether the hydration of cosolute exclusion would be the major contribution to observed effects.

1.9 Effects of Crowding on Biological Systems

1.9.1 Crowding Effects on Proteins

The Osmophobic Effect

Nucleic acid structures are generally destabilized by osmolytes but stabilized by macrosolutes (22). Proteins however, can experience positive effects in the presence of osmolytes as a result of the osmophobic effect (97). For example, researchers reported up to 3.2-fold increases in the dimerization of α-chymotripsin in the presence of sucrose, glucose and raffinose (98). The protein backbone solvation is at the center of the osmophobic effect (97). Peptide groups are the dominant group in proteins and they contribute to the total change in exposed surface area of the protein. These peptide groups prefer water as a solvent rather than osmolytes. Thus, the peptides are
osmophobic, and they tend to fold in the presence of osmolytes in order to reduce the backbones exposure to the osmolytes (97). The side chains of a protein are osmophilic, however they do not contribute to the stabilization effect that is observed with osmolytes. Thus in the presence of osmolytes, refolding of α-chymotrypsin is accelerated and the unfolding velocity is reduced.

**Macrosolute Effects on Proteins**

Proteins can be stabilized by macrosolutes via steric repulsion. For example, the thermal denaturation of hen egg-white lysozyme was increased by 2.5 °C in the presence of 10% (w/v) dextran 35,000 (99). Additionally, a conformational change resulted in a more compact form of the protein. The conformational change is in agreement with the idea that crowding will favor the more compact form of a structure. Protein folding can also be affected by crowding. A urea induced unfolded form of ribonuclease A was placed in crowded conditions created by PEG 20,000 and Ficoll 70. In 35% (w/v) solutions of PEG and Ficoll, the unfolded protein converted to the native and compact conformation (100). Addition of PEG 200 to the unfolded protein also resulted in folding, but not in a compact structure.

Protein association is concentration dependent. For example, in order for fibrinogen to form a dimer, sufficiently high concentrations are required. However, in the presence of 10% BSA, dilute fibrinogen formed a dimer (101). This association
reaction was linked to an increase in activity coefficient to 10 for fibrinogen in the presence of BSA, versus an activity coefficient of 1 in the dilute solution.

The state of a protein, such as ordered versus disordered, is important in understanding the physiological relevance of a protein because the state of a protein can alter its activity. The primary component of lewy bodies is α-synuclein, a natively disordered protein (102). Parkinson’s disease is histologically characterized by the formation of lewy body aggregates. Although α-synuclein is disordered, it is a collapsed structure which occupies less volume than the ordered form of the protein. At 35°C, α-synuclein undergoes a conformational change where the hydrodynamic radius is increased, and 100 residues on the N-terminus form a secondary structure (102). In the presence of 30% BSA, the disordered but compact form of the protein was stabilized, and no conformational change was noted at 35°C (102).

1.9.2 Crowding Effects on Nucleic Acids

One of the first to observe the effects of crowding on nucleic acids was Arthur Kornberg. After studying in vitro DNA replication for almost a decade, Kornberg found success in his experiments when he added high concentrations of PEG. The presence of PEG resulted in a crowded environment which stabilized the interactions between proteins and the origin of replication (103). Recently, Kornberg made mention of the importance of correcting for cellular concentrations by naming macromolecular
crowding as one of the ten rules that must be respected in experiments (104). In more recent years, crowding has been utilized to explore numerous nucleic acid reactions.

**Duplex and Triplex DNA**

Duplex and triplex DNA can be stabilized in the presence of cosolutes. For example, the $T_m$ of DNA homopolymers, poly(dAT) duplexes, were increased 3-10 °C in the presence of PEGs and dextrans (105). Evidence of duplex stabilization in a crowded environment was demonstrated by Goobes and Minsky, who investigated crowding effects on short (17-20 base pairs) poly (dAT) duplexes and triplexes with and without inosine mismatches (48,106). The duplexes and triplexes were studied in the presence of macrosolutes (PEG of molecular weight 1000, 3350, 6000 and 8000), and osmolytes (ethylene glycol and PEG 200) at concentrations ranging from 4.5-15 % (w/v). The $T_m$’s of the duplexes in 4.5% osmolyte were identical to the dilute environment, indicating that the osmolyte had no effect. However, when the concentration of the osmolytes was increased to 15%, the $T_m$ of the duplexes decreased by 5 °C (106). Conversely, crowding induced stabilization was observed for all macrosolutes at 15% concentration with ~4 °C increase in $T_m$ for all duplexes, including the inosine mutants (106). The crowding effects noted for triplex structures were even greater with increases in $T_m$ of up to 16 °C in 15% PEG 8000. None of the osmolytes had a destabilizing effect on the triplexes, however, 15% PEG 200 resulted in an 8 °C increase in $T_m$ (106).
The differences in the effects of duplex versus triplex are believed to be due to the difference in their hydration. Triplex DNA releases water upon formation whereas duplex DNA takes up water. Also, Goobes and Minsky concluded that the nature and/or number of their mismatch had no effect on the crowding mediated effects observed (106). Goobes and Minsky also concluded that the effects of macrosolutes were excluded volume mediated, while osmolyte effect were water activity mediated (106).

The stability of duplex and triplex DNA was also studied by Spink and Chaires. Poly(dA)⋅poly(dT) duplexes and poly(dA)⋅poly(dA)⋅poly(dT) triplexes were studied in macrosolutes (PEG of molecular weight 1000, 3400 and 8000) and osmolytes (ethylene glycol, glycerol and PEG 400) at concentrations of 5-20 % (w/v) (61). The osmolytes resulted in a depression of melting temperature for both triplex and duplex at 20% concentrations. However, at 20% concentrations the macrosolutes yielded an increase in $T_m$. In other published reports, Chaires and Spink studied the melting enthalpy of poly(dGdC)⋅poly(dCdG) duplexes in the presence of various osmolytes (ethylene glycol, glycerol, sucrose, urea, betaine, PEG 200) and in the presence of a macrosolute (PEG 1450) (107). All of the osmolytes except betaine and sucrose resulted in a decrease of the melting enthalpy of the duplexes. The decrease was attributed to accumulation of the osmolytes on the surface of the DNA which resulted in a small increase of melting enthalpy. In the presence of macrosolute, both duplexes exhibited increases in $T_m$. The results of Chaires and Spink emphasize the importance
of cosolute size (107). Recent investigations into accumulation of cosolutes on DNA surfaces confirm the findings of Spink and Chaires (70). It was proven that osmolytes accumulate on the surface of single stranded DNA after duplex melting. This is the generally accepted view of how osmolytes modulate water activity resulting in a destabilization of nucleic acids regardless of sequence.

Sugimoto has also studied the effects of crowding on duplex DNA. He has investigated how the nucleotide length modulates crowding effects for duplex thermal denaturation (108). Using 8, 17, and 30 base pair oligonucleotides, Sugimoto concluded that oligonucleotide length and crowding effects are directly related. Sugimoto studied the duplexes in the presence of macrosolutes (PEG 1000 and 8000) and osmolytes (ethylene glycol and PEG 200) at 20% (w/v) concentrations. His findings showed that only one of the 30-mer oligonucleotide exhibited an increase in Tm of 0.8 °C in the presence of PEG 8000. However, Sugimoto employs concentrations of 1M NaCl, which have been demonstrated to result in decreases in the crowding effects (106). Also, the stabilization Sugimoto observes is only of 0.8 °C, it is a negligible amount compared to Chaires and Spink who reported that 20 base pair duplexes have exhibited up to 5 °C increase in Tm in the presence of 20% PEG 8K (22). Under the same conditions, Chaires and Spink have reported a 5 °C increase in Tm for a 200 base pair duplex, which contradicts the statements made by Sugimoto that crowding effects and duplex length are related.
In other investigations, Sugimoto studied the kinetics of duplex DNA formation in the presence of macrosolute (PEG 8000) and osmolytes (PEG 200 and 600) at concentrations of 20% (w/v) in 1M NaCl (109). The presence of the osmolytes resulted in a decrease of the rate constant of formation by up to 1.7-fold. The macrosolute showed a slight increase in the rate constant of formation of 1.1-fold. It was concluded that the osmolytes raised the activation energy barrier for the formation of duplex while the macrosolute had no effect on the activation barrier. These results are at odds with the work presented in this thesis as will be discussed in future chapters.

Finally, reactions of \( \lambda \) DNA (~24,000 base pairs) were studied by Louie and Serwer (110). The bimolecular association and cyclization of \( \lambda \) DNA were studied under crowding conditions in the presence of osmolyte (PEG 200) and macrosolute (PEGs of molecular weight greater than 1540). Louie and Serwer found that the osmolytes shifted both the association and cyclization reactions towards dissociation as a result of reduced water activity (110). Conversely, the macrosolutes shifted both reactions towards association.

**Quadruplex DNA**

Quadruplex DNA has been implicated in the inhibition of telomerase activity, a quality that provides for the potential use of quadruplexes as anticancer agents (111). A sequence which is capable of forming a quadruplex or an anti-parallel hairpin looped duplex was studied in the presence of an osmolyte (40% PEG 200) (112). Stabilization
of an anti-parallel G-quadruplex and destabilization of the duplex was noted. In contrast to duplex DNA, quadruplex DNA actually dehydrates (releases water) upon formation. Quadruplex DNA is similar to triplex DNA in that an osmolyte can result in an enhancement of stability. Researchers concluded that when water activity is decreased by cosolutes, for structures where dehydration is a characteristic, a favorable effect can be observed (112).

Another interesting quadruplex crowding discovery involved the conversion of a parallel G-quadruplex to an anti-parallel one. Sugimoto’s research group studied the effects of crowding on C-rich and G-rich DNA which were capable of forming an I-motif and anti-parallel quadruplex respectively in dilute solutions (16). In addition, they also studied the effects of crowding on the duplex formed by the 1:1 mixture of the G-rich and C-rich DNA. In crowded environments, created by the osmolyte, 60% PEG 300, the G-rich DNA formed a parallel quadruplex, while the C-rich DNA maintained the conformation it had in the dilute solution. Interestingly, under crowded conditions, the 1:1 mixture of G and C-rich DNA did not form a duplex, but formed the parallel quadruplex and the I-motif (16). It is important to note that the prevention of duplex formation once again indicates that duplexes are affected negatively by osmolytes.
References


Chapter 2– Macrosolute Effects on a Bimolecular System with Single Mismatches

2.1 Introduction

Single nucleotide polymorphisms (SNPs) are variations that occur in the sequence of DNA. SNP classification is given to a variation when a single nucleotide in the genome is different between at least 1% of a species (1). SNPs in coding sequences of genes can predispose humans to diseases such as cancer, or result in changes in response to pathogens (2). Although the majority of SNPs are found in non-coding regions, they are biologically relevant due to their health implications. Detection of known SNPs can be carried out with the use of mutant and wild-type probes which hybridize to the sequence of interest (3). The presence of a mismatched base between the probe and target indicates a possible SNP.

There are two types of mismatched bases, transversion and transition (4). In a transversion mismatch, two purine or two pyrimidines are paired together. Alternatively, a transition mismatch occurs when a purine is paired with the incorrect pyrimidine. The presence of a mismatch is disruptive to a helix, but the extent of disruption is dependent on the identity of the mismatched base pair (sections 1.2-1.3). A mismatched base pair stability scale was developed by Ke and Wartell (5,6). According to the scale, a CC base pair mismatch is the most destabilizing to a duplex while the least destabilizing is a GG base pair. Both the CC and GG base pairs are transversion mismatches. A single-base mismatch in a duplex can be difficult to detect...
because the difference in energy between a complementary duplex and a single-
mismatched duplex is small, typically only 3–6 kcal mol\(^{-1}\) (7).

A wild-type (WT) duplex and two single mismatched base containing duplexes were used to determine if differential stabilization of one duplex over another was possible in a crowded environment. The CC and GG base pairs were incorporated into a duplex and were selected because of their locations within the mismatched stability scale. A difference in stabilization could be the result of differences in the structure and stability of the resulting mismatched duplexes. The effects of crowding in a competitive environment were also studied to see if a duplex could be stabilized against strand invasion. The kinetics of duplex formation on the mismatched base duplexes were also studied in a crowded environment to see if the formation of one duplex over another was enhanced. In order to monitor the thermodynamic and kinetic reactions, fluorescence resonance energy transfer (FRET) was used.

**The BRCA1-185 System**

The sequences for the bimolecular thermodynamic and kinetic studies were selected on the basis of their biologically relevant and heterogeneous sequence. The use of homopolymers such as poly(dAT), which were utilized by Minsky and Goobes to investigate crowding effects on duplexes, do not always reflect natural systems (8,9). Furthermore, with homopolymers certain structural problems may arise, such as A-tract
bending (10,11). Generating a sequence is also not ideal as it is less relevant and may lead to the formation of undesired secondary structures.

The BRCA1-185 sequence was selected for the bimolecular studies because of its implication in breast and ovarian cancer (Scheme 2.1). The BRCA1-185 sequence is used in genetic testing, including allele-specific oligonucleotide hybridization assays (12). A WT duplex and two single mismatched duplexes (CC and GG) were derived from the BRCA1-185 sequence. The sequences used are not known to form any secondary structures according to Integrated DNA Technologies Inc. OligoAnalyzer 3.1™.

The BRCA1-185 sequence is located on chromosome 17q and is the sequence where several mutations linked to cancer are found (13). These mutations are believed to be partly responsible for the early onset of certain cancers (14). There are three common mutations found within the Ashkenazi population; two of the mutations are on BRCA1 and the other is on BRCA2. The two BRCA1 mutations are 5382insC, insertion of a cytosine at the 5382 position, and 185delAG in which an adenine (A) and a guanine (G) at position 185 are deleted (15). The mutation on BRCA2 is 8174delT, the deletion of a thymine (T) at position 6174 (16). It is estimated that >2% of Ashkenazi Jewish population carry at least one of the mutations on BRCA-1 (15).

Men and women who have the 185delAG mutation have a 56% lifetime risk of developing breast cancer (14). In addition there is a 16% lifetime risk of ovarian and prostate cancer for women and men, respectively (16). Within the Ashkenazi
population, people with one or more of the mutations have an average age of cancer
onset that is almost 10 years earlier than individuals with no mutations (17). The two
mutations on BRCA-1 occur at a rate of 1/300 in non-Ashkenazi women, but occur at a
rate of 1/40 for the Ashkenazi women (15). As a result, people of Ashkenazi Jewish
descent are encouraged to seek genetic testing to determine if they are carriers of the
mutations.

<table>
<thead>
<tr>
<th></th>
<th>5’ AA TCT TAG AGT GTC CCA 3’</th>
<th>TAMRA</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>3’ TT AGA ATC TCA CAG GGT 5’</td>
<td>WT (A) TAMRA</td>
</tr>
<tr>
<td></td>
<td>5’ AA TCT TAG AGT CTC CCA 3’</td>
<td>TAMRA</td>
</tr>
<tr>
<td>B</td>
<td>3’ TT AGA ATC TCA CAG GGT 5’</td>
<td>CC (A) TAMRA</td>
</tr>
<tr>
<td></td>
<td>5’ AA TCT TAG AGT GTC CCA 3’</td>
<td>TAMRA</td>
</tr>
<tr>
<td>C</td>
<td>3’ TT AGA ATC TCA CAG GGT 5’</td>
<td>GG (B) TAMRA</td>
</tr>
</tbody>
</table>

Scheme 2.1 – The BRCA1-185 duplexes. (A) The WT duplex (B) the CC duplex and
(C) the GG duplex. The underlined bolded bases represent the mismatched bases. The
donor fluorophores, 5’6-carboxyflourescein (FAM), and the acceptor fluorophore,
tetramethylrhodamine (TAMRA) were placed on the same end of the duplex to
maximize the proximity of the fluorophores for detection purposes.

BRCA1-185 Studies Monitored via FRET

Two fluorophores, 5’6-carboxyflourescein (FAM) and tetramethylrhodamine
(TAMRA) were placed on the same end of the duplex. The proximity of the two
fluorophores resulted in quenching of the FAM intensity by TAMRA as a result of
fluorescence resonance energy transfer (FRET). In FRET an interaction takes place
between the electronic excited states of two dye molecules, in which excitation is
transferred from a donor molecule to an acceptor molecule without emission of a
photon (18). The non-radiative interaction between the donor (D) and acceptor (A) fluorophores occurs through intermolecular long range dipole-dipole coupling (19). The quenching of the FAM intensity was used to monitor the progress of the BRCA1-185 thermal denaturation, competition studies, and kinetic reactions.

Some common applications of FRET include measuring the distances between macromolecules both in vitro and in vivo. The structure and stability of nucleic acids have also been studied extensively with FRET. For example, FRET was employed to investigate the dissociation and association of helices, as well as the formation and dissociation of multi-branched DNA junctions (20-22). Additionally, the FRET reaction between molecular beacons and targets have been used to measure expression levels of mRNA in cells (23,24).

In order for FRET to take place, certain requirements must be satisfied. The first is the proper selection of fluorophores; the absorption spectrum of the acceptor must overlap the emission spectra of the donor (25). Second, the quantum yield of the donor and the absorption coefficient of the acceptor must be sufficiently high (26). FRET is also strongly dependent on the proximity of the fluorophores, a relationship which was derived by Förster in the 1940’s (27). Förster determined that the energy transfer rate ($k_{ET}$) between donor and acceptor, scales as the sixth power of their separation (R).

$$k_{ET} = \frac{1}{\tau_D} \left( \frac{R_0}{R} \right)^6 \quad \text{(Equation 2.1)}$$
In equation 2.1, the fluorescence lifetime of the donor in the absence of the acceptor is represented by $\tau_D$, while $R_0$ is the critical distance at which $k_{ET}$ equals the intrinsic decay rate. It is estimated that in order for FRET to take place, the fluorophores must be between 10-100 Å apart (18). The $R_0$ is defined in terms of the spectral properties of the fluorophores (20).

$$R_0^6 = 8.8 \times 10^{-28} \Phi_D \kappa^2 n^{-4} J(\lambda) \quad \text{(Equation 2.2)}$$

In equation 2.2, $\Phi_D$ is the fluorescence quantum yield of the donor, the orientational dependence of the energy transfer is $\kappa^2$, the refractive index of the medium that separates the donor and acceptor is $n$, and $J(\lambda)$ is the spectral overlap between the donor emission and acceptor absorption.

The energy transfer efficiency, $E$ is the fraction of the excited donor molecules that are decaying when FRET takes place (equation 2.3).

$$E = \frac{1}{1 + \left(\frac{R}{R_0}\right)^6} \quad \text{(Equation 2.3)}$$

As the distance between the fluorophores increases, the energy transfer efficiency decreases. Thus, to maximize the energy transfer rate, the fluorophores for BRCA1-185 were placed on the same end of the duplex (Scheme 2.1). FAM has an excitation maximum at 494 nm and an emission maximum at 520 nm. TAMRA has excitation and emission maximum at 559 nm and 583 nm respectively. The $R_0$ value for the FRET pair used in the BRCA1-185 studies is approximately 58 Å (19). The reporter
wavelength for the thermodynamic, kinetic and competition studies was 520 nm, the
FAM peak emission. The loss or gain of FAM intensity was used to monitor the
progress of the reactions. For the thermal denaturation studies, as the strands melted
and moved apart the FAM intensity increased as a result of the loss of FRET.
Conversely, for the kinetic studies, as the hybridization progressed and the strands
moved closer together, the FAM intensity decreased due to the occurrence of FRET.

2.2 Materials and Methods

2.2.1 Materials

Chemicals

The salts, MgCl$_2$, NaCl, KCl, Na$_2$HPO$_4$ and NaH$_2$PO$_4$ were purchased from
Sigma-Aldrich. The polymers, poly(ethylene glycol) (PEG) of average molecular
weight 3400 (3.4K), 8000 (8K) and 15000-20000 (15K), Ficoll type 70 and glycerol
were also purchased from Sigma-Aldrich. Bovine serum albumin (BSA) was purchased
from Fisher Scientific. The BSA was further purified by two rounds of dialysis against
water. The polymers were dissolved in phosphate buffered saline (PBS). The cosolute
concentrations are reported as weight per volume (w/v).
Oligonucleotides

The labeled and unlabeled oligonucleotides were obtained from Integrated DNA Technologies Inc (Coralville, IA). The oligonucleotides were dissolved in PBS (pH 7.4, 137 mM NaCl, 2.6 mM KCl, 10 mM Na$_2$HPO$_4$ and 1.8 mM NaH$_2$PO$_4$). The concentrations of the oligonucleotides were determined by ultra-violet (UV) absorbance on a Beckman spectrophotometer (model DU-640), using the extinction coefficients at 260 nm (Table 2.1). Complementary oligonucleotides were mixed in a 1:1 concentration ratio for the formation of duplex. The oligonucleotides were annealed by heating at $95^\circ$C for 4 minutes after which they were cooled to room temperature over a period of two hours.

<table>
<thead>
<tr>
<th></th>
<th>$\varepsilon_{260\text{ nm}}$ Labeled (M$^{-1}$ cm$^{-1}$)</th>
<th>$\varepsilon_{260\text{ nm}}$ Unlabeled (M$^{-1}$ cm$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>WT (A)</td>
<td>166,900</td>
<td></td>
</tr>
<tr>
<td>WT (A) FAM</td>
<td>187,860</td>
<td></td>
</tr>
<tr>
<td>WT (B)</td>
<td>168,800</td>
<td></td>
</tr>
<tr>
<td>WT (B) TMR</td>
<td>201,800</td>
<td></td>
</tr>
<tr>
<td>GG (A)</td>
<td>170,600</td>
<td></td>
</tr>
<tr>
<td>GG (A) FAM</td>
<td>191,560</td>
<td></td>
</tr>
<tr>
<td>CC (B)</td>
<td>165,300</td>
<td></td>
</tr>
<tr>
<td>CC (B) TMR</td>
<td>205,300</td>
<td></td>
</tr>
</tbody>
</table>

Table 2.1 – The extinction coefficients (260 nm) for the BRCA1-185 sequences. The extinction coefficients were obtained from IDT Inc. and were based on the salt concentrations of PBS buffer.
2.2.2 Measurements

Thermal Denaturation of the BRCA1-185 Duplexes

Melt profiles were obtained in a dilute environment (PBS) and in crowded environments created by the presence of macrosolutes (PEG 3.4K, 8K, 15K and Ficoll 70) and an osmolyte (glycerol). The macrosolute and osmolyte concentrations used were 10 and 20% for each. Thermal denaturation of the duplexes was monitored via FRET as the quenching of FAM intensity at 520 nm on a Photon Technologies Incorporated fluorimeter (model QM-2001-4). All melts were conducted in a 1.0 cm path length quartz cuvette. An external bath was used to control the temperature of the cell holder. Samples were heated in intervals of 5-10 °C, where a 5 °C temperature interval was used in the transition region of the melt. The equilibration time for the cuvette at each temperature interval was determined with the use of an external thermal couple. Once equilibration at a temperature interval was complete, the samples were excited at 494 nm and emission spectra were collected from 510 nm to 600 nm at a scanning rate of one nm minute$^{-1}$. The excitation and emission slit widths were 3 and 6 nm respectively. A 480-499 nm band-pass filter obtained from Chroma Technology Corporation (Rochingham, VT) was used in the excitation beam to reduce light scattering. Excitation correction was used to correct for fluctuations in the lamp intensity. All scans were background corrected against buffer (PBS) or crowded solution. The thermal denaturation profiles were created by graphing the intensity at 520 nm ($I_{520}$) versus temperature. The concentration used was 10 nM per strand for all of the duplexes. The thermal denaturation studies were faced with experimental
limitations. The equipment available did not allow for automatic temperature control and constant manual adjustment of temperature resulted in low data density. A second limitation was the total experiment duration. However, values reported were based on the average of a minimum of three individually fit trials per experiment.

**Kinetics of Hybridization of the BRCA1-185 Duplexes**

Hybridization profiles were obtained in a dilute environment (PBS) and in crowded environments simulated by the presence of macrosolutes (20% PEG 3.4K, 8K, Ficoll type 70 or BSA) and an osmolyte (20% glycerol). The progress of the hybridization reactions was monitored via the quenching of FAM intensity by FRET on a Photon Technologies Incorporated fluorimeter (model QM-2001-4). All kinetics reactions were conducted in a 1.0 cm path length quartz cuvette. An external water bath was used to maintain the temperature of the cuvette holder at 25 °C. Complementary oligonucleotides were allowed to individually equilibrate at 25 °C for 10 minutes. Upon completion of the equilibration period, the samples were mixed and data collection was initiated. A time-based emission scan was collected at 520 nm (FAM peak) with excitation at 494 nm at 25 °C. An automatic shutter was employed to prevent photo-bleaching of the dyes. Scans were collected by opening the shutter for five seconds and collecting data at a rate of 0.2 points per second. Following each five second collection window was a 15 second period during which no data was collected and the shutter was closed. The total collection times, 30 or 60 minutes, were dependent on the sequences
and conditions. The excitation and emission slit widths were 3 and 6 nm respectively. A 480-499 nm band-pass filter obtained from Chroma Technology Corporation (Rockingham, VT) was used in the excitation beam to reduce light scattering. Excitation correction was used to correct for fluctuations in the lamp intensity. All scans were background corrected against buffer (PBS) or crowded solution. The kinetics of hybridization profiles were created by graphing the normalized intensity at 520 nm (I_{520}) versus time. In kinetics of hybridization studies, the concentrations used were 2.5, 5 or 10 nM per strand. The concentration of the strands was determined by the cosolute used. The values reported are based on the fit of an average data set obtained from three or more individual trials.

**Competition Studies on BRCA1-185 WT**

In the competition assays, an unlabeled competitor strand was added to a previously annealed WT duplex which was labeled with FAM and TAMRA. A series of samples were created with increasing concentration of unlabeled competitor and a fixed concentration of WT duplex. Upon addition of the competitor strand to the duplex, the samples were annealed in a Perkin Elmer Cetus thermal cycler (model TC1). The samples were heated and maintained at 95 °C for 4 minutes after which the temperature was cooled to 4 °C at a rate of 0.5 °C minute^{-1}. The samples were maintained at 4 °C for 30 minutes and then heated to room temperature and allowed to equilibrate for one hour. The reactions were conducted in a dilute environment (PBS
with added 5 mM MgCl₂) and in crowded environments created by the presence of a macrosolute (20% PEG 8K in PBS with 5 mM MgCl₂).

The quenching of the FAM intensity by FRET was used to monitor the progress of the strand competition reactions. Measurements were conducted on a Photon Technologies Incorporated fluorimeter (model QM-2001-4). For all of the competition reactions, a 1.0 cm path length quartz cuvette was used. An external water bath maintained the temperature of the cell holder at 25 °C. After the samples were equilibrated in the cuvette, an emission scan was initiated. Spectra were obtained by exciting the sample at 494 nm and collecting data from 510 nm to 600 nm at a scanning rate of 1 nm minute⁻¹. The excitation and emission slit widths were 3 and 6 nm respectively. A 480-499 nm band pass filter obtained from Chroma Technology Corporation (Rockingham, VT) was used in the excitation beam to reduce light scattering. Excitation correction was used to correct for fluctuations in the lamp intensity. All scans were background corrected against buffer (PBS) or crowded solution. The strand invasion profiles were created by graphing the intensity at 520 nm (I₅₂₀) versus the concentration of competitor strand. The concentration of the WT duplex strands was 2.5 nM for the WT(A)TAMRA strand, and 5 nM for the WT(B)FAM strand. The ratio of the labeled duplex strands was 1:2 (FAM strand: TAMRA strand) to ensure that all of FAM labeled strands were annealed. The unlabeled competitor strand concentration ranged from 68 µM to 2 nM. The values
reported are based on the fit of an average data set obtained from three or more individual trials.

2.2.3 Data Analysis

Analysis of the BRCA1-185 Melt Profiles

The melt profiles of the duplexes were obtained by plotting the normalized intensity at 520 nm (I_{520}) versus temperature in Kelvin (K). The melt profiles were fit with an expression for two-state melting of a non-self complementary sequence using Kaleidagraph 3.6. The association of the BRCA1-185, non-self-complementary sequences, can be expressed as:

\[ A + B \rightleftharpoons C \]  \hspace{1cm} (Equation 2.4)

The molecularity (n) of the reaction is the number of strands that comprise the n-mer structure. Thus, for the BRCA1-185 duplex, n is two. The equilibrium constant, \( K_{eq}(T) \), for the association described above can be expressed as:

\[
K_{eq}(T) = \frac{[AB]}{[A][B]} = \frac{C_T}{n} \frac{\alpha}{\alpha(1-\alpha)^n} = \frac{2\alpha}{C_T(1-\alpha)^2} \]  \hspace{1cm} (Equation 2.5)

where \( C_T \) represents the total strand concentration and \( \alpha \) is the fraction of double-stranded DNA (28,29). Mathematica 6.0 was used to solve equation 2.5 in terms of \( \alpha \). The negative form of the root was real in the range of the experimental data.
\[ \alpha(T) = \frac{1 + C_T K_{eq}(T) - \sqrt{1 + 2C_T K_{eq}(T)}}{C_T K_{eq}(T)} \]  \hspace{1cm} (Equation 2.6)

The van’t Hoff expression for \( K_{eq}(T) \) is:

\[ K_{eq}(T) = \exp \left( -\frac{\Delta G}{RT} \right) = \exp \left( -\frac{\Delta H}{RT} + \frac{\Delta S}{R} \right) \]  \hspace{1cm} (Equation 2.7)

where \( \Delta G \) is the change in free energy, \( \Delta H \) is the change in enthalpy, \( \Delta S \) refers to the change in entropy and \( R \) is the gas constant, 1.987 cal mol\(^{-1}\) K\(^{-1}\) (28,29). Substitution of the van’t Hoff expression for \( K_{eq}(T) \) into equation 2.6 yields an expression for \( \alpha \) in terms of \( \Delta H \) and \( \Delta S \).

\[ \alpha(T) = \frac{1 + C_T \exp \left( -\frac{\Delta H}{RT} + \frac{\Delta S}{R} \right) - \sqrt{1 + 2C_T \exp \left( -\frac{\Delta H}{RT} + \frac{\Delta S}{R} \right)}}{C_T \exp \left( -\frac{\Delta H}{RT} + \frac{\Delta S}{R} \right)} \]  \hspace{1cm} (Equation 2.8)

The intensity at 520, \( I_{520} \), at a given temperature can be defined by:

\[ I_{520}(T) = \alpha [I_{520}(T_{\text{min}}) + m_{ds}(T - T_{\text{min}})] + (1 - \alpha)[I_{520}(T_{\text{max}}) - m_{ss}(T_{\text{max}} - T)] \]  \hspace{1cm} (Equation 2.9)

where \( I_{520}(T_{\text{min}}) \) is the average measured intensity at the temperature minimum \( (T_{\text{min}}) \) and \( I_{520}(T_{\text{max}}) \) is the average measured intensity at the temperature maximum \( (T_{\text{max}}) \) (29). The \( m_{ds} \) and \( m_{ss} \) are the average slopes for the first and last three temperature points respectively. The \( m_{ds} \) represents the temperature dependence of the intensity of the double-stranded (ds) DNA and \( m_{ss} \) represents the temperature dependence of the
intensity of the single-stranded (ss) DNA (29). This procedure corrected for the temperature dependence of the upper and lower baselines.

The \( \alpha \) expression in equation 2.8 was substituted into equation 2.9, and used to analyze the melting profiles for the BRCA1-185 duplexes. The enthalpy and entropy terms were obtained from the curve fit analysis and were used to calculate the melting point, \( T_m \). The general equation for \( T_m \) (equation 2.10) can be applied to a non-self-complementary system of any molecularity (28).

\[
\frac{1}{T_m} = \frac{(n-1)R}{\Delta H} \ln C_T + \left[ \frac{\Delta S - (n-1)R \ln 2n}{\Delta H} \right] \quad (\text{Equation 2.10})
\]

The \( T_m \) expression for the BRCA1-185 (\( n=2 \)) system can be defined as:

\[
T_m = \frac{\Delta H}{\Delta S + R \ln \left( \frac{C_T}{4} \right)} \quad (\text{Equation 2.11})
\]

The Gibbs free energy (\( \Delta G \)) was calculated at 37 °C because it is physiologically relevant (29). Equation 2.12 was also used to calculate the \( K_{eq(37C)} \) for the Kirkwood-Buff analysis.

\[
\Delta G = \Delta H - T \Delta S = -R T \ln K_{eq} \quad (\text{Equation 2.12})
\]

The melt profiles were individually fit and \( \Delta H, \Delta S, \Delta G_{37C} \) and \( T_m \) values were obtained. The melt profiles were individually fit to obtain \( \Delta H, \Delta S \) and calculate \( \Delta G_{37C} \) and \( T_m \) values. The reported \( \Delta H, \Delta S, \Delta G_{37C} \) and \( T_m \) are average values based on a minimum of 3 trials. The error for parameters was calculated as standard error of the mean.
Analysis of the BRCA1-185 Hybridization Profiles

The kinetics of duplex formation is a second order reaction where $A + B \rightarrow AB$ and the rate law can be expressed as

$$r = k[A][B] \quad \text{(Equation 2.13)}$$

where $[A]=[B]$. As a result of the equimolar ratio, equation 2.13 can be simplified using $[S]$ which represents the single strands that comprise the duplex.

$$r = k[S]^2 \quad \text{(Equation 2.14)}$$

Integration of the rate law, and separation of variables yields an expression which describes $[S]$ at any time ($t$) where at $t=0$, the $[S]_t$ is equal to the sum of $[A]_0$ and $[B]_0$.

$$[S]_t = \frac{[S]_0}{1 + kt[S]_0} \quad \text{(Equation 2.15)}$$

The $[S]_t$ is then used to express $I_{520}(t)$, the intensity at 520 nm at time $t$.

$$I_{520}(t) = f_1 \left( \frac{[S]_0}{1 + kt[S]_0} \right) + f_2 \left( [S]_{\text{total}} - \frac{[S]_0}{1 + kt[S]_0} \right) \quad \text{(Equation 2.16)}$$

In equation 2.16, $f_1$ is the signal contribution from the single strands and $f_2$ is the signal contribution from the duplex, $t$ is time in seconds, and $k$ is the rate constant in $M^{-1} s^{-1}$. Prior to curve fitting, the data was normalized using the following equation:

$$I_{520}(t) = \frac{I_{520}(t) - I_{520}(\text{min})}{I_{520}(\text{max}) - I_{520}(\text{min})} \quad \text{(Equation 2.17)}$$

where $I_{520}(\text{min})$ and $I_{520}(\text{max})$ refer to the minimum and maximum intensity at 520 nm respectively. Three or more normalized data sets were averaged to obtain the rate.
constants of hybridization. The hybridization profiles were curve-fit using Kaleidagraph 3.6.

**Analysis of the BRCA1-185 WT Competition Studies**

Analysis of the competition studies was done with Mathematica 6.0. The equation used to fit the competition data was adapted from Lin and Riggs (30).

\[
\theta_0 = \frac{\theta T (1 - \theta_0)}{K_{AD} \left(1 + \frac{X_T}{K_{AX}}\right) + D_T (1 - \theta_0)} \quad \text{(Equation 2.18)}
\]

In equation 2.18, the fraction of competitor duplex (\(\theta\)) is described in terms of the concentrations of the donor (\(D_T\)), acceptor (\(A_T\)), and competitor strands (\(X_T\)). The donor strand was WT(B) FAM and the acceptor strand was WT(A) TAMRA. The \(K_{AD}\) represents the dissociation constant (\(K_d\)) of the WT duplex. The \(K_{AD}\) values were calculated from the WT hybridization parameters in PBS and 20% PEG 8K using equation 2.12. The \(K_{AX}\) represents \(K_d\) of the duplex formed with the competitor strand and one of the strands from the WT duplex. Using Mathematica 6.0, equation 2.18 was solved in terms of \(\theta\) and the positive root yields real values for the range of the data.

\[
\theta = \frac{-\left(D_T + K_{AD} + A_T + X_T \frac{K_{AD}}{K_{AX}}\right) + \sqrt{\left(D_T + K_{AD} + A_T + X_T \frac{K_{AD}}{K_{AX}}\right)^2 - 4D_T A_T}}{2D_T} \quad \text{(Equation 2.19)}
\]

Two terms were used to account for the range and minimum of the data:
\[ \theta = f_1 \left( -\left( \frac{D_T + K_{AD} + A_T + X_T \left( \frac{K_{AD}}{K_{AX}} \right)}{2D_T} \right) + \sqrt{\frac{\left( D_T + K_{AD} + A_T + X_T \left( \frac{K_{AD}}{K_{AX}} \right) \right)^2 - 4D_T A_T}{2D_T}} \right) + f_2 \]  

(Equation 2.20)

where \( f_1 \) is the range and \( f_2 \) is the minimum.

2.3 Results

2.3.1 Macrosolute Effects on the Thermodynamics of the BRCA1-185 Duplexes

The BRCA1-185 duplexes (WT, CC and GG) were selected to study the effects of crowding on the thermal stability of duplexes with single mismatches. The duplexes were heated and the loss of secondary structure was monitored by the changes in FAM quenching at 520 nm. As the temperature was increased the duplex was disrupted and the strands grew farther apart. As the average distance between the fluorophores increased, the intensity of the FAM peaks rose in accordance with the single strand intensity value. Endpoint values for the melts were obtained by comparing the isolated single strand intensity at the same concentration as the duplex.

Figure 2.1a depicts the relative intensities of 10 nM WT single strands and a 10 nM WT duplex. The FAM labeled strand (WT (B) FAM) was excited at 494 nm and exhibited an emission peak at 520 nm with an intensity of about 310,000 counts per second (cps). The TAMRA labeled strand was excited at both 547 and 494 nm.
Excitation of the TAMRA labeled strand (WT(A)TAMRA) at 547 nm resulted in an emission peak at 580 nm with an intensity of approximately 350,000 (cps). When the same TAMRA labeled strand is excited at 494 nm, the emission peak at 580 nm has a lower intensity of 40,000 cps. Upon duplex formation, the intensity of the FAM peak drops (from 300,000 to 100,000 cps), while the intensity of the TAMRA peak increases (from 40,000 to 60,000 cps). The decrease in FAM peak is indicative of quenching that may be a result of the proximity between the fluorophores.

Figure 2.1b shows a representative melt of a CC duplex in the presence of 20% PEG 3.4K. The duplex was heated and as the temperature increased the intensity of the FAM increased indicating loss of FAM quenching. The representative melt profile for CC in the presence of 20% PEG 3.4K (Figure 2.1c) was obtained by graphing the $I_{520}(T)$ versus temperature. Using the thermodynamic expression (equation 2.9) the data was fit to extrapolate and calculate the thermodynamic parameters. Melting of a duplex labeled with FAM and TAMRA resulted in an increase of approximately 450,000 cps in FAM intensity, a 7.2-fold increase (Figure 2.1b). When a duplex labeled only with FAM is melted, the resulting change in FAM intensity was a 90,000 cps decrease, a 1.4-fold decrease (Figure 2.1d). The differences in the change in signal between the two duplexes (double labeled and single labeled) indicated that the quenching of FAM intensity was a result of FRET due to the proximity of the TAMRA fluorophore. The thermodynamics and kinetic studies of the duplexes were monitored as the change in FAM intensity due to the larger change in signal.
The duplexes were studied under dilute and crowded conditions. The dilute conditions for all duplexes were created by using a simple buffer, PBS. The crowded environments were created by the presence of 10 or 20% macrsolute (PEG 3.4K, 8K, 15K or Ficoll type 70). Additionally, the duplexes were studied in the presence of an osmolyte, 20% glycerol. The use of PBS as a control allowed for the comparison of a crowded environment versus a simple dilute buffer. Hence, the functional significance of the cosolute presence could be observed. Figures 2.2, 2.3 and 2.4 contain the individual average graphs for the WT, CC and GG duplexes respectively. Table 2.2 displays the thermodynamic parameters obtained from the fit (ΔH and ΔS) and the calculated parameters (T_m and ΔG_{37C}) for the WT, CC and GG duplexes in dilute and crowded conditions.
Figure 2.1 – Representative data from the BRCA1-185 thermal denaturation studies. (A) Sample data of 10 nM WT single strands and 10 nM duplex. The WT(B)FAM strand (green), WT(A)TAMRA strand (grey) and the WT duplex (black) were excited at 494 nm. Additionally, the WT(A)TAMRA strand (pink) was also excited at its excitation maximum of 547 nm. The intensity of the FAM peak decreases when going from the single stranded to the duplex form and the intensity of the TAMRA peak increases. (B) Sample experimental data from the thermal denaturation of a CC duplex in the presence of 20% PEG 3.4K. As the temperature increases, the intensity of the FAM peak increases indicating loss of duplex structure. (C) The sample data from panel B and other CC melts in 20% PEG 3.4K were averaged and normalized. Error bars shown are based on a minimum of 3 trials. The melt profile was fit (blue) to obtain ΔH and ΔS. (D) Thermal denaturation of a WT duplex in PBS where only the FAM fluorophore was present. As the temperature increased, the intensity of FAM decreased.
Figure 2.2 – Thermal denaturation of the WT duplex in PBS and various macrosolutes. (A) PBS (B) 10% PEG 3.4K (C) 20% PEG 3.4K (D) 10% PEG 8K (E) 20% PEG 8K. The solid lines represent the curve-fit analysis of the melt profiles. A minimum of 3 trials were conducted per condition.
Figure 2.3 – Thermal denaturation of the WT duplex in the presence of various macrosolutes and an osmolyte. (A) 10% PEG 15K (B) 20% PEG 15K (C) 10% Ficoll (D) 20% Ficoll (E) 20% glycerol. The solid lines represent the curve-fit analysis of the melt profiles. A minimum of 3 trials were conducted per condition.
Figure 2.4 – Thermal denaturation of the CC duplex in PBS and various macrosolutes. (A) PBS (B) 10% PEG 3.4K (C) 20% PEG 3.4 K (D) 10% PEG 8K (E) 20% PEG 8K. The solid lines represent the curve-fit analysis of the melt profiles. A minimum of 3 trials were conducted per condition.
Figure 2.5 – Thermal denaturation of the CC duplex in the presence of various macrosolutes and an osmolyte. (A) 10% PEG 15K (B) 20% PEG 15K (C) 10% Ficoll (D) 20% Ficoll (E) 20% glycerol. The solid lines represent the curve-fit analysis of the melt profiles. A minimum of 3 trials were conducted per condition.
Figure 2.6 – Thermal denaturation of the GG duplex in the presence of various cosolutes. (A) PBS (B) 20% PEG 3.4K (C) 20% PEG 8K (D) 10% PEG 8K (E) 20% PEG 8K. The solid lines represent the curve-fit analysis of the melt profiles. A minimum of 3 trials were conducted per condition.
<table>
<thead>
<tr>
<th></th>
<th>$\Delta H$ (kcal mol$^{-1}$)</th>
<th>$\Delta S$ (cal mol$^{-1}$K$^{-1}$)</th>
<th>$\Delta G_{37^\circ C}$ (kcal mol$^{-1}$)</th>
<th>$T_m$ (°C)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>WT Duplex</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PBS</td>
<td>-108.6 ± 1.8</td>
<td>-297.2 ± 5.9</td>
<td>-16.4 ± 0.1</td>
<td>50.9 ± 0.4</td>
</tr>
<tr>
<td>10% PEG 3.4K</td>
<td>-116.5 ± 0.6</td>
<td>-317.1 ± 1.9</td>
<td>-17.9 ± 0.1</td>
<td>54.1 ± 0.1</td>
</tr>
<tr>
<td>20% PEG 3.4K</td>
<td>-117.8 ± 0.4</td>
<td>-318.1 ± 1.3</td>
<td>-19.1 ± 0.1</td>
<td>57.6 ± 0.1</td>
</tr>
<tr>
<td>10% PEG 8K</td>
<td>-116.1 ± 0.7</td>
<td>-316.8 ± 1.6</td>
<td>-17.9 ± 0.2</td>
<td>54.1 ± 0.4</td>
</tr>
<tr>
<td>20% PEG 8K</td>
<td>-115.9 ± 0.9</td>
<td>-315.1 ± 2.4</td>
<td>-18.2 ± 0.1</td>
<td>55.1 ± 0.2</td>
</tr>
<tr>
<td>10% PEG 15K</td>
<td>-116.2 ± 0.7</td>
<td>-316.7 ± 2.2</td>
<td>-18.0 ± 0.1</td>
<td>54.5 ± 0.2</td>
</tr>
<tr>
<td>20% PEG 15K</td>
<td>-116.8 ± 0.7</td>
<td>-317.4 ± 2.3</td>
<td>-18.3 ± 0.1</td>
<td>55.4 ± 0.2</td>
</tr>
<tr>
<td>10% Ficoll 70</td>
<td>-114.7 ± 0.6</td>
<td>-312.5 ± 1.6</td>
<td>-17.8 ± 0.1</td>
<td>54.2 ± 0.1</td>
</tr>
<tr>
<td>20% Ficoll 70</td>
<td>-115.3 ± 0.3</td>
<td>-315.0 ± 1.0</td>
<td>-17.6 ± 0.1</td>
<td>53.6 ± 0.2</td>
</tr>
<tr>
<td>20% Glycerol</td>
<td>-89.2 ± 0.4</td>
<td>-242.7 ± 1.5</td>
<td>-13.9 ± 0.1</td>
<td>44.5 ± 0.3</td>
</tr>
<tr>
<td><strong>CC Duplex</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PBS</td>
<td>-104.0 ± 1.3</td>
<td>-299.2 ± 4.8</td>
<td>-11.2 ± 0.1</td>
<td>35.4 ± 0.3</td>
</tr>
<tr>
<td>10% PEG 3.4K</td>
<td>-105.7 ± 0.7</td>
<td>-303.5 ± 2.1</td>
<td>-11.6 ± 0.1</td>
<td>36.5 ± 0.2</td>
</tr>
<tr>
<td>20% PEG 3.4K</td>
<td>-106.0 ± 1.4</td>
<td>-302.6 ± 4.1</td>
<td>-12.2 ± 0.2</td>
<td>38.1 ± 0.5</td>
</tr>
<tr>
<td>10% PEG 8K</td>
<td>-106.5 ± 1.1</td>
<td>-305.4 ± 3.4</td>
<td>-11.8 ± 0.1</td>
<td>37.0 ± 0.2</td>
</tr>
<tr>
<td>20% PEG 8K</td>
<td>-106.6 ± 0.3</td>
<td>-305.4 ± 0.9</td>
<td>-11.9 ± 0.1</td>
<td>37.3 ± 0.3</td>
</tr>
<tr>
<td>10% PEG 15K</td>
<td>-106.4 ± 0.3</td>
<td>-303.8 ± 0.7</td>
<td>-12.1 ± 0.1</td>
<td>38.1 ± 0.4</td>
</tr>
<tr>
<td>20% PEG 15K</td>
<td>-107.9 ± 0.3</td>
<td>-308.3 ± 1.0</td>
<td>-12.3 ± 0.1</td>
<td>38.5 ± 0.4</td>
</tr>
<tr>
<td>10% Ficoll 70</td>
<td>-106.0 ± 0.7</td>
<td>-305.3 ± 2.4</td>
<td>-11.3 ± 0.1</td>
<td>35.5 ± 0.1</td>
</tr>
<tr>
<td>20% Ficoll 70</td>
<td>-104.0 ± 0.7</td>
<td>-299.3 ± 2.6</td>
<td>-11.2 ± 0.1</td>
<td>35.3 ± 0.2</td>
</tr>
<tr>
<td>20% Glycerol</td>
<td>-91.3 ± 0.7</td>
<td>-265.1 ± 2.7</td>
<td>-8.9 ± 0.4</td>
<td>28.2 ± 0.8</td>
</tr>
<tr>
<td><strong>GG Duplex</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PBS</td>
<td>-106.8 ± 0.7</td>
<td>-302.9 ± 2.1</td>
<td>-12.9 ± 0.1</td>
<td>40.2 ± 0.2</td>
</tr>
<tr>
<td>20% PEG 3.4K</td>
<td>-109.9 ± 0.1</td>
<td>-310.0 ± 1.0</td>
<td>-13.8 ± 0.1</td>
<td>42.7 ± 0.1</td>
</tr>
<tr>
<td>20% PEG 8K</td>
<td>-112.7 ± 0.5</td>
<td>-317.7 ± 1.9</td>
<td>-14.2 ± 0.1</td>
<td>43.7 ± 0.2</td>
</tr>
<tr>
<td>20% Glycerol</td>
<td>-101.6 ± 0.3</td>
<td>-291.7 ± 0.9</td>
<td>-11.1 ± 0.1</td>
<td>34.9 ± 0.1</td>
</tr>
</tbody>
</table>

Table 2.2 – The thermodynamic parameters of hybridization for the WT, CC and GG BRCA1-185 duplexes in dilute and crowded conditions. All values reported are based on the average of a minimum of 3 trials. Reported error values were tabulated as the standard error of the mean.
Table 2.3 – The changes in $T_m$ and $\Delta G_{37^\circ C}$ for the WT, CC and GG duplexes in the presence of cosolutes. The $T_m$ and $\Delta G_{37^\circ C}$ values were calculated as the difference of the values in crowded versus dilute environment (PBS) from Table 2.2. Reported error values were tabulated using propagation of error.
The temperature at which half of the duplex structure is lost is the melt temperature (T_m). In the dilute environment the T_m of the WT, CC and GG duplexes was 50.9, 35.4 and 40.2 °C respectively (Table 2.2). Although there is only a single base difference between the three duplexes, their T_m’s in PBS vary greatly. The presence of the CC base pair results in a 15.5 °C decrease in T_m (PBS) as compared to the WT duplex. An 10.7 °C decrease in T_m is noted for the GG duplex. The relative differences in T_m are in agreement with the mismatched base stability scales of Santalucia and others (6,31,32). Of all the possible mismatched base pairs, CC and GG are the most and least disruptive to a duplex respectively.

The difference between the T_m of the duplexes in the crowded and dilute environment, ΔT_m, describes the effect of a crowded environment on a particular duplex (Table 2.3). In order to compare the three duplexes to each other, the change in the free energy in the crowded versus the dilute environment was used, ΔΔG_{37°C}. The ΔΔG_{37°C} can be used as a standard for comparing various structures that melt at different temperatures. In the presence of macrosolute, the values of ΔT_m were positive for all duplexes and the values of ΔΔG_{37°C} were negative for all duplexes (Table 2.3). Hence, the three duplexes were stabilized by the presence of all the macrosolutes.

The ΔT_m range for the WT duplex, in the presence of macrosolutes, was 2.7 to 6.7 °C. The largest crowding-mediated effect was noted for the WT duplex in the presence of 20% PEG 3.4K with a 6.7 °C increase in T_m. The ΔΔG_{37°C} for the WT duplex in 20% PEG 3.4K was -2.7 kcal mol^{-1}. The CC duplex exhibited a ΔT_m range of
1.1 to 3.1 °C in the presence of macrosolutes. The greatest stabilization observed for the CC duplex was in the presence of 20% PEG 15K in which the $T_m$ increased by 3.1 °C. The $\Delta G_{37^\circ C}$ was -1.1 kcal mol$^{-1}$ for the CC duplex in the presence of 20% PEG 15K. The presence of the macrosolutes resulted in a $\Delta T_m$ range of 2.5 to 3.3 °C for the GG duplex. The presence of 20% PEG 8K stabilized the GG duplex to the greatest extent. The $\Delta G_{37^\circ C}$ of -1.3 kcal mol$^{-1}$ for the GG duplex in 20% PEG 8K. Thus, in the presence of macrosolute, the WT duplex was stabilized to a greater extent than either one of the mismatched duplex. The only osmolyte used in the thermodynamic studies was glycerol. All of the duplexes were destabilized in the presence of glycerol as indicated by the positive $\Delta G_{37^\circ C}$ and negative $\Delta T_m$ values (Table 2.3).

The results of the thermodynamic studies were used to construct plots of $\Delta G_{37^\circ C}$ versus temperature (Figure 2.7). The $\Delta G$ values were calculated (from 280 to 370 K) for all of the duplexes in PBS, 20% PEG 3.4K, 8K and glycerol, by using their respective curve-fit obtained $\Delta H$ and $\Delta S$ values. The plots demonstrate the temperature dependence of the free energy. The slopes of the lines are equal to the $\Delta S$ values determined by the curve fit. For all of the duplexes, the slope was the smallest in the presence of the osmolyte, glycerol. The greatest slope for the WT and CC duplexes was in the presence of 20% PEG 3.4K. While, the GG duplex displayed the greatest slope was for the 20% PEG 8K. The separation between the PBS and macrosolute lines as the temperature decreases is the greatest for the WT duplex. These plots demonstrate
the greater extent of stabilization of the WT duplex over the mismatch containing duplexes.

![Graphs showing free energy vs. temperature for WT, CC, and GG duplexes in various cosolutes.](image)

**Figure 2.7 - The relationship between the free energy and the temperature of the duplexes in cosolutes.** (A) The WT duplex (B) the CC duplex and (C) the GG duplex in the presence of PBS (black), 20% PEG 3.4K (green), 8K (blue) and glycerol (red).
2.3.2 Macrosolute Effects on the Kinetics of Duplex Formation of BRCA1-185

In addition to thermodynamic consequences, crowding can also affect the kinetics of duplex formation by increasing the rate of the forward reaction. The addition of cosolute to a system leads to an increase in the viscosity which is expected to decrease the rate of association (33). However, an increase in the chemical activity opposes viscosity effects and an increase in association may occur (34). The rate of duplex formation was measured for the WT, CC and GG duplexes via FAM quenching by FRET. As the hybridization reaction progressed, the intensity of the FAM peak decreased as FRET occurred due to the increasing proximity of the strands. Hybridization profiles were collected in dilute conditions (PBS) and in crowded environments created by macrosolutes (20% PEG 3.4K, 8K, Ficoll 70 or bovine serum albumin). Bovine serum albumin (BSA) was used as a macrosolute in order to employ a more physiological cosolute. BSA is a globular protein with a molecular weight of 66,000 Da. Additionally, crowded conditions were created with an osmolyte, 20% glycerol. Cosolute concentration and selection were based on the effects noted for the BRCA1-185 thermodynamic studies (section 2.3.1).
The results of the hybridization kinetic experiments are listed in Table 2.4. The rate constant of hybridization in PBS for the WT, CC and GG duplexes was 7.3, 13.7 and $8.2 \times 10^5$ M$^{-1}$s$^{-1}$ respectively. The values reported here are in agreement with the typical rate constant of duplex formation of $10^5$ to $10^7$ M$^{-1}$s$^{-1}$ (35,36). In the presence of all the macrosolutes (PEG 3.4K, 8K, Ficoll 70 and BSA), an increase in the rate constant of formation was observed as compared to the dilute buffer (PBS). As the macrosolute size increased, the enhancement of the rates of hybridization increased.

The greatest rate constant enhancement for all three duplexes was in the presence of the physiological macrosolute, BSA. The rate constant of duplex formation increased 4.4-fold in the presence of BSA for WT. In the presence of BSA, the CC and GG exhibited a 3.5 and 3.6-fold increase in rate constant respectively. The presence of the osmolyte, glycerol, resulted in a decrease in the rate constant of hybridization for all duplexes. The effect of glycerol on the duplexes was similar for all three duplexes.

<table>
<thead>
<tr>
<th></th>
<th>WT K (10$^5$ M$^{-1}$ s$^{-1}$)</th>
<th>CC K (10$^5$ M$^{-1}$ s$^{-1}$)</th>
<th>GG K (10$^5$ M$^{-1}$ s$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1X PBS</td>
<td>$7.3 \pm 1.1$</td>
<td>$13.7 \pm 0.3$</td>
<td>$8.2 \pm 0.4$</td>
</tr>
<tr>
<td>20% Ficoll 70</td>
<td>$10.0 \pm 1.2$</td>
<td>$14.7 \pm 0.3$</td>
<td>$8.1 \pm 0.4$</td>
</tr>
<tr>
<td>20% PEG 3.4K</td>
<td>$18.0 \pm 0.2$</td>
<td>$22.9 \pm 0.1$</td>
<td>$22.1 \pm 0.1$</td>
</tr>
<tr>
<td>20% PEG 8K</td>
<td>$23.3 \pm 0.2$</td>
<td>$35.9 \pm 0.4$</td>
<td>$33.1 \pm 0.6$</td>
</tr>
<tr>
<td>20% BSA</td>
<td>$32.2 \pm 0.3$</td>
<td>$48.3 \pm 0.8$</td>
<td>$29.6 \pm 0.4$</td>
</tr>
<tr>
<td>20% Glycerol</td>
<td>$3.8 \pm 0.9$</td>
<td>$6.0 \pm 1.3$</td>
<td>$4.5 \pm 0.1$</td>
</tr>
</tbody>
</table>

Table 2.4 – The rate constants of hybridization for the WT, CC and GG duplexes. All rates are based on 3 or more trials which were averaged and then fit. Reported error values are the standard error of predication obtained from the fit.
2.3.3 Macrosolute Effects on the WT BRCA1-185 Duplex in a Competitive Environment

Competition experiments were conducted in order to exploit the differential stabilization of the WT duplex over both single mismatched duplexes. These experiments aimed at probing the effects of crowding on the stability of a duplex in a competitive environment. The unlabeled competitor strands (WT(B) and CC (B)) were titrated into samples containing a fixed concentration of previously annealed labeled WT duplex. The competition reactions were monitored via FAM quenching by FRET. As the concentration of competitor strand increased, and the competitor strand invaded the WT duplex, an increase in the intensity of the FAM peak was observed. The dilute environment was created with PBS. The crowded conditions were created by the presence of 20% PEG 8K, a macrosolute which enhanced the stability of BRCA1-185 duplexes against thermal denaturation (section 2.3.1).

The competition of the WT(B) strands in the dilute and crowded environments yielded identical curves indicating no effect as a result of the macrosolute (Figure 2.8a). However, when a mismatch containing competitor strand was introduced, CC(B), a difference was observed between the crowded and dilute environments (Figure 2.8a). The competition data was analyzed and $K_{AX}$ values were obtained from the fit. $K_{AX}$ is the $K_d$ of the newly formed duplex created by the competitor strand and one of the WT duplex strands. The competition of the CC(B) strand in PBS resulted in a $K_{AX}$ of $1.82 \times 10^{-10}$ M. In the presence of 20% PEG 8K, $K_{AX}$ was $1.67 \times 10^{-11}$ M. Hence in the
presence of macrosolute, the WT duplex was stabilized against strand displacement by competitor strand.

![Figure 2.8 – Competition studies in dilute and crowded environments.](image)

(A) The WT(B) strand and (B) the CC(B) strand were titrated into in a fixed concentration of labeled WT Duplex. The closed circles represent the competition in PBS and the open circles represent the competition in the presence of PEG 8K. The data points and error bars shown are based on the average of 3 or more trials. The solid lines represent the curve fit using equation 2.21.

### 2.4 Discussion

#### 2.4.1 Differential Stabilization of a WT Duplex Over Both Single Base Mismatched Duplexes

Excluded volume theory states that in a crowded environment, the more compact form of a macromolecule will be favored. Crowding favors compaction because it results in an increase in the available volume to the cosolutes. In the case of B-form DNA, the covolume of the duplex is less than that of the single strands (37). Hence, in a crowded environment, a duplex can be stabilized. The thermodynamic
studies were successful in demonstrating crowding mediated enhancement of duplex stability in the presence of macrosolutes. As the concentration of a specific macrosolute increased, the observed effects also increased. However, the relationship between cosolute concentration and observed effect is not necessarily linear. The BRCA1-185 stability studies also highlighted the importance of appropriate cosolute size selection. All of the duplexes were destabilized in the presence of the osmolyte, but stabilized by the macrosolutes.

The thermodynamic studies on the BRCA1-185 duplexes confirmed the relative stabilities of the mismatched bases. Both of the mismatches incorporated into the duplexes were homo-transversion mismatches that were selected due to their contrasting effects on the duplex stability. In PBS, the duplex containing the CC mismatch was destabilized significantly, while the duplex containing the GG mismatch was destabilized to a lesser extent. The destabilizing effect of the mismatches to the duplex can be correlated with their significant differences in $\Delta G_{37C}$. As compared to the $\Delta G_{37C}$ for WT in PBS, the $\Delta G_{37C}$ of the CC and GG duplexes in PBS increased (became less negative) by 5.2 and 3.5 kcal mol$^{-1}$ respectively.

The WT duplex was differentially stabilized over both mismatched duplexes as demonstrated by the $\Delta \Delta G_{37C}$ values, which compared the difference between the $\Delta G_{37C}$ of a duplex in crowded versus dilute environment (Table 2.3). The greatest stabilization for the WT duplex was in 20% PEG 3.4K with a $\Delta \Delta G_{37C}$ of -2.4 kcal mol$^{-1}$. The CC and GG duplexes had $\Delta \Delta G_{37C}$ respective values of -1.0 and -0.9 kcal mol$^{-1}$ in the
presence of 20% PEG 3.4K. Hence, the two mismatched duplexes were stabilized to similar extents. Comparison of the change in $\Delta\Delta G_{37C}$ between the duplexes (in 20% PEG 3.4K) demonstrates that the WT duplex was stabilized differentially over both single mismatched base duplexes at an average of -1.7 kcal mol$^{-1}$.

The differential crowding-mediated enhancement of the WT duplex by macrosolutes is contrary to the proposal by Abraham Minsky, who stated that crowding-mediated stabilization is independent of the number and nature of the mismatched bases (38). Minsky’s conclusion is based on his work on oligonucleotides in the presence of macrosolutes and osmolytes. Although the oligonucleotides that Minsky studied were of similar length to the BRCA1-185 strands, the duplexes that he studied were poly(dA)·poly(dT) with inosine-adenine (IA) and GA mismatches. In addition to being non-canonical, IA base pairs, are not very destabilizing to a helix (6,39). Also, GA base pairs are quite stable (section 1.4). Additionally, Minsky utilized experiments with cosolute concentrations of only up to 15% (w/v). The total increase in $T_m$ reported by Minsky was 3-5 °C for both wild-type and mismatched duplexes in 0.1 M NaCl and only a 1°C increase in $T_m$ in 1M NaCl experiments. The $\Delta T_m$ range for our WT duplex in macrosolute was similar to those reported by Minsky. However, the macrosolute induced increase in $T_m$ was much less for the CC and GG duplex.

The data presented here also does not correlate with Sugimoto and colleagues, who report a 0.8 °C decrease in stabilization for a 17-mer oligonucleotide in 20% PEG
8K and a 0.3 °C increase in stabilization for a 30-mer (40). However, the salt concentrations that Sugimoto utilized were 1M NaCl, and as shown by Minsky, crowding effects decrease significantly as ionic strength increases. Thus, Sugimoto’s findings are not consistent with our conditions. However, the macrosolute effects reported here are consistent with observations by Chaires and Spink, who reported that macrosolutes result in stabilization of duplexes, while osmolytes led to destabilization against thermal denaturation (41,42). Chaires and Spink noted that both macrosolutes and osmolytes alter water activity. However, the volume exclusion effects by macrosolutes are greater than the changes in hydration. Conversely, osmolytes mainly mediate their crowding effects through changes in water activity.

**Kirkwood Buff Theory Interpretation of Crowding Effects**

Kirkwood-Buff theory of solutions can be used to interpret the crowding-mediated stabilizations that were observed for the BRCA1-185 duplexes (section 1.8) (43-47). The BRCA1-185 system is described as consisting of water (i=1), a macromolecule (i=2) and cosolute (i=3) (section 1.8). Using these components, the change in the preferential hydration parameter ($\Delta \nu_{21}$) can be described as the effect that water (in terms of chemical potential $\mu_1$) has on a macromolecule at a constant pressure (P), temperature (T), and molality of the macromolecule ($m_2$) (43,45,48). The $\Delta \nu_{21}$ can
also be described as the change in the number of cosolutes ($\Delta N_{23}$) and water molecules ($\Delta N_{21}$) distributed around the surface of the macromolecule, in terms of molality (44).

\[
\Delta \nu_{21} = RT \left( \frac{\partial \ln K}{\partial \mu_1} \right)_{T,P,n_2} = \Delta N_{21} - \frac{m_1}{m_3} \Delta N_{23} \quad \text{(Equation 2.21)}
\]

\[
\Delta N_{21} = n_1 N_A \Delta G_{21} = n_1 N_A (G_{21}^B - G_{21}^A) \quad \text{(Equation 2.22)}
\]

\[
G_{21}^\sigma = \int d\bar{r} [g_{21}^\sigma (\bar{r}) - 1] \quad \text{(Equation 2.23)}
\]

The $\Delta N_{21}$ and $\Delta N_{23}$ are directly correlated to the radial distribution of the water and cosolute around the duplex respectively (equations 2.22 and 2.23). The $\Delta N_{21}$ value is positive for duplex formation since water is taken up in this process. Conversely, $\Delta N_{21}$ is negative when a duplex is disrupted since approximately 4 waters are released per base pair (41,42). Using equation 2.21, the thermodynamic parameters reported in Table 2.2 and published values for changes in chemical activity of water in PEG 8K and glycerol, the $\Delta \nu_{21}$ was evaluated for the duplexes in these cosolutes (40).
Table 2.5 – The change in the preferential hydration ($\Delta v_{21}$) upon addition of cosolutes. The change in preferential hydration parameter was calculated using equation 2.21, published values for change in water activity upon cosolute addition and the thermodynamic parameters from Table 2.2.

<table>
<thead>
<tr>
<th></th>
<th>$\Delta T_m$ (°C)</th>
<th>$\Delta \ln K_{eq}$ (mol$^{-1}$)</th>
<th>$\Delta v_{21}$ (mol mol$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>WT 20% PEG 8K</td>
<td>3.6</td>
<td>0.582</td>
<td>Negative</td>
</tr>
<tr>
<td>WT 20% Glycerol</td>
<td>-5.6</td>
<td>-0.737</td>
<td>Positive</td>
</tr>
<tr>
<td>CC 20% PEG 8K</td>
<td>1.7</td>
<td>0.232</td>
<td>Negative</td>
</tr>
<tr>
<td>CC 20% Glycerol</td>
<td>-7.5</td>
<td>-0.931</td>
<td>Positive</td>
</tr>
<tr>
<td>GG 20% PEG 8K</td>
<td>3.1</td>
<td>0.387</td>
<td>Negative</td>
</tr>
<tr>
<td>GG 20% Glycerol</td>
<td>-5.5</td>
<td>-0.659</td>
<td>Positive</td>
</tr>
</tbody>
</table>

All of the duplexes had a negative $\Delta v_{21}$ in the presence of macrosolute. Using equation 2.21, and the fact that $\Delta N_{21}$ is positive for hybridization reactions, certain generalizations can be made concerning the BRCA1-185 duplexes. The negative $\Delta v_{21}$ values for the duplexes in the presence of 20% PEG 8K indicate that the $\Delta N_{23}$ must be positive and larger than the $\Delta N_{21}$. Thus, the general assumption of $\Delta N_{23}$ for duplexes in the presence of 20% PEG 8K is that the distribution of PEG is greater around the duplex than the single strands and those changes in cosolute exclusion are greater than changes in hydration (Figure 2.9). We conclude that the stabilization effects, on the BRCA1-185 duplexes by macrosolutes, are dominated by the changes in macrosolute distribution and exclusion. However, contributions may arise from changes in hydration which are not as great as the excluded volume effects.
In the presence of osmolyte, all of the duplexes had a positive $\Delta v_{21}$ which correlates to a possibly positive $\Delta N_{23}$ which is smaller than the $\Delta N_{21}$. Our evaluations are consistent with published reports since glycerol has been shown to significantly alter water activity (9,38,40). Therefore, although both $\Delta N_{21}$ and $\Delta N_{23}$ contribute to the $\Delta v_{21}$ value, the major contribution to $\Delta v_{21}$ in the case of osmolytes is $\Delta N_{21}$. We conclude that the observed osmolyte effects on BRCA1-185 were mediated mostly by the change in hydration which led to the destabilization of the duplexes.

**Figure 2.9 - A schematic representation of the distribution of macrosolutes and waters upon duplex formation and dissociation.** The larger circles represent a unit of volume composed of water (i=1, blue spheres), DNA (i=2) and 20% macrosolute (i=3, red spheres). Bulk solvent properties are observed outside of the dashed circle. The inner circle represents the immediate vicinity surround DNA, the local domain. In environments crowded with macrosolutes, upon hybridization the $\Delta N_{21}$ is positive, but smaller than the $\Delta N_{23}$. Hence, the macrosolute distribution around the vicinity of the duplex is greater than around the single strands. For clarity, only a representative number of water molecules are shown.
2.4.2 Enhancement of Duplex Formation in the Presence of Macrosolutes

The formation of the WT, CC and GG duplexes were studied in the presence of various cosolutes. The macrosolute presence resulted in an increase in the forward rate of the duplex formation reaction. Up to a 4.4-fold increase in the rate constant of duplex formation was observed. The enhancement is due to a decrease in the activity energy barrier. It has been shown that duplex formation is dependent on nucleation of the first few bases, which determines the energy barrier (50). In the thermodynamic studies of the BRCA1-185 duplexes, differential stabilization of the WT over both single mismatched base duplexes was noted. Although the rate constant of formation was enhanced, differential enhancement of one duplex over another was not observed.

The rate constant enhancement observed was dependent on the identity of the cosolute and its size. As the size of our macrosolutes increased, an increase in the rate constant of hybridization was noted. This trend was much more noticeable for the kinetic studies than for the thermodynamic studies of BRCA1-185. The smaller macrosolutes, PEG 3.4 and 8K did not yield the same extent of enhancement of the rate constant as the larger BSA macrosolute did. Furthermore, the importance of the cosolute shape is demonstrated by Ficoll 70 and BSA. Although their molecular weights are similar, their shapes are different. The Ficoll is a rod-like cosolute, but the shape of the BSA is more globular. The presence of Ficoll resulted in minimal enhancement in the rate constant of duplex formation. However, the duplexes in the presence of BSA resulted in the greatest enhancement of the rate constant for all three duplexes.
2.4.3 Crowding Results in the Stabilization of WT BRCA1-185 Against Competition

The competition experiments investigated if differential stabilization of the WT duplex over a single mismatched duplex could be observed in a competitive environment at room temperature. The titration of the CC(B) strand into the WT duplex resulted in the displacement of one strand of the WT duplex and subsequent formation of a duplex which contained a CC mismatched base pair. The competition studies contained an internal control. When the competitor strand was the same identity as one of the WT duplex strands, no change was noted in dilute versus crowded environment. This occurred because when the WT(B) strand was the competitor, the resulting new duplex formed was a WT duplex where one strand is unlabeled. As expected, the results of the competition of WT(B) in PBS and 20% PEG 8K yielded similar values for $K_{AX}$.

Differences in $K_{AX}$ were noted in PEG versus PBS when the CC(B) strand was the competitor. When the fit $K_{AX}$ in PBS is converted into $\Delta G_{25C}$, the value is -14 kcal mol$^{-1}$, a value which is close to the $\Delta G_{25C}$ obtained from the thermodynamic studies of CC duplex in PBS (-15.3 kcal mol$^{-1}$). Conversion of the fit $K_{AX}$ in 20% PEG 8K into $\Delta G_{25C}$ yields -15.3 kcal mol$^{-1}$, a value similar to the thermodynamic $\Delta G_{25C}$ of the CC duplex in PEG 8K (-15.9 kcal mol$^{-1}$). Hence, the comparison of $\Delta \Delta G_{25C}$ from the competition experiments in PBS and PEG 8K is -1.3 kcal mol$^{-1}$ and indicates the extent stabilization of the WT against strand displacement.
The curve-fit expression used for the competition analysis contained $K_{AD}$ values in PBS and PEG for the WT duplex (which were obtained from the thermodynamic studies). As a result, the $\Delta\Delta G_{25^\circ C}$ from the competition experiments is a comparison of the $\Delta\Delta G_{25^\circ C}$ for WT versus CC duplex. The ratio of the CC(B) $K_{AX}$ in PBS and PEG is directly related to the differences in free energy of the two duplexes in dilute versus crowded environments. The thermodynamic studies of the WT, CC and GG BRCA1-185 duplexes demonstrated differential stabilization of the WT duplex over both single base mismatched duplexes (sections 2.3.1 and 2.4.1). From the thermodynamic studies, the $\Delta\Delta G_{25^\circ C}$ for the WT duplex in 20% PEG 8K compared to PBS is $-1.7$ kcal mol$^{-1}$.

Also, the $\Delta\Delta G_{25^\circ C}$ for the CC duplex in 20% PEG 8K compared to PBS is $-0.6$ kcal mol$^{-1}$. The resulting differential stabilization of the WT over CC duplex is $-1.1$ kcal mol$^{-1}$ at 25 °C. From the competition studies, we learn that the WT duplex was stabilized against strand displacement by the CC(B) strand, by $-1.3$ kcal mol$^{-1}$. A comparison of the thermodynamic and competition experiment $\Delta\Delta G_{25^\circ C}$ values leads to the conclusion that differential stabilization of the WT duplex over the CC duplex can also be observed in the context of a competitive environment.

The $K_{AX}$ values reported here are only approximations since the competition system was limited. Total displacement of the WT duplex was not possible due to the limitations in the concentration of the competitor strand for two reasons. First, at high concentrations of DNA, precipitation can occur in the presence of PEG. Second, the concentration of the original stock of competitor strand was limiting. To correct for the
limitations due to the concentration of competitor, the end point for the titration of the CC(B) competitor strand was approximated from the control experiments of incoming WT(B). These end points were incorporated as constraints on the curve-fitting of the CC(B) competitor data.

2.5 Conclusions

The BRCA1-185 system is a bimolecular system (n=2) that was studied under crowded conditions to determine thermodynamic effects of crowding against thermal denaturation and strand displacement. The effects of crowding on the kinetics of duplex formation were also studied. The thermodynamic data presented here demonstrates differential stabilization of a WT duplex over two single-base mismatched duplexes in the presence of macrosolutes. An additional aspect of interest is that the single mismatched bases were on opposite sides of the mismatch stability scale. The GG mismatch is the least destabilizing to a duplex, while CC is the most destabilizing. Yet, in thermodynamic assays, both mismatched duplexes were stabilized to similar extents, and both were stabilized to a lesser extent than the WT duplex. All three duplexes were destabilized by the presence of an osmolyte, glycerol, which demonstrated the importance of size considerations when selecting a cosolute. The competition assays demonstrated that the differential stabilization of the WT over the CC duplex could also be observed at room temperature.
The thermodynamic effects were interpreted using Kirkwood-Buff solution theory which demonstrated that the stabilization induced by macrosolutes is due to cosolute exclusion. However, the destabilizing effect induced by osmolytes is due to hydration changes. The differential crowding-mediated enhancement by neutral polymers observed was contrary to the proposal that stabilization is independent of the number and nature of the mismatched bases. In general, the least destabilizing single-mismatched duplex, GG, was stabilized to a lesser extent than the most destabilizing single-mismatch duplex. Interestingly, we also saw a larger range of stabilization than previously reported. In contrast to the thermodynamics studies, the kinetic studies did not demonstrate differential enhancement of formation of any of the three duplexes. Instead, the rate constant of formation for all three duplexes was enhanced in the presence of macrosolutes and decreased in the presence of an osmolyte. These results indicate a possible method for increasing the speed and specificity of single base mutant detection in hybridization assays.
References


Chapter 3 – Macrosolute Effects on Bimolecular Systems with Single Base Bulges

3.1 Introduction

The insertion or deletion of a base(s) in double helical DNA results in a bulge base(s) which can be incorporated into the helix (intrahelical) or can loop out and lie within the groove of the duplex (extrahelical) (1). Intrahelical bulges squeeze in between existing bases and can cause frame-shift mutations and stretching of the phosphodiester backbone of the opposite strand (2). Extrahelical bulges however, do not participate in base pairing and the integrity of the helix sequence is otherwise preserved (3). The extrahelical conformation of a bulge base can result in increased hydration of a bulge duplex as compared to the parent, fully complementary, duplex (4).

A bulge base(s) can arise from errors during replication or from recombination between strands (5). DNA bulges can range from one to ten bases and are repaired by the methyl-directed mismatch repair pathway (6). RNA bulge regions have been implicated as recognition sites for proteins (7). Feedback regulation and tertiary folding of RNA have also been linked to bulge regions (8,9). The biological relevance of bulge bases highlights the necessity to understand how they are influenced by their naturally crowded environment (10,11). Previously, in our study of the BRCA1-185 duplexes in the presence of macrosolutes, a -1.6 kcal mol$^{-1}$ differential stabilization of a WT duplex over two mismatched duplexes was observed (chapter 2). However, in the BRCA1-185 kinetics of hybridization studies, the presence of the macrosolutes did not result in preferential acceleration of the association of one duplex over another. Here we
present a study on single base bulge duplexes to determine if differential stabilization is noted for one duplex over another. Also, the kinetics of hybridization studies will focus on investigating if the bulge containing duplexes exhibit preferential kinetic effects in crowded environments.

The 5382 System

A biologically relevant and heterogeneous sequence was selected for the single base bulge duplex thermodynamic and kinetic studies. All of the sequences used in the studies were derived from BRCA1-5382insC (Scheme 3.1), a sequence that has been linked to the predisposition of breast and ovarian cancer within the Ashkenazi Jewish population (12). BRCA1-5382insC is the insertion of a cytosine at position 5382 on BRCA1 which is located on chromosome 17q and is one of three common mutations within the Ashkenazi population (13,14). The second common mutation, also located on BRCA1, is 185delAG, the deletion of an adenine and guanine at position 185 (15). The BRCA1-185 sequence was used to study the effects of crowding on duplexes containing single mismatched base pairs (chapter 2). The third mutation is located on BRCA2, and is 8174delT, in which a thymine is deleted at position 8174 (16). These mutations are linked to an average onset of breast cancer that is almost ten years younger for carriers versus non-carriers (15). As a result, the Ashkenazi are encouraged to seek genetic testing to determine if they are carriers of one of the mutations (15).
Scheme 3.1 – The 5382 duplexes. (A) The unlabeled 5382-WT duplex (B) the unlabeled 5382-X bulge duplex (C) the labeled 5382-WT duplex and (D) the labeled 5382-C duplex. The unlabeled duplexes in A and B were used in the thermodynamic studies which were monitored via CD. The duplexes in C and D were labeled with a donor fluorophore (FAM) and an acceptor fluorophore (HEX) in order to monitor the kinetics of duplex formation via the quenching of FAM intensity by FRET. The fluorophores were placed on the same end of the duplex in order to maximize the proximity of the fluorophores for detection purposes.

The 5382 duplexes were used to probe the crowding effects on duplexes with a single bulge base (Scheme 3.1). The wild-type duplex (5382-WT) was the parent duplex which did not contain a bulge base. There is no stability scale for the bulge bases because their extent of helix disruption is dependent on the identity of both the bulge base and its neighboring bases (17). For this reason, all four canonical bases were incorporated as single bulges in the thermodynamic studies. The bulge duplexes are collectively referred to as 5382-X. The 5382 position was the location of base insertion of an adenine (A), cytosine (C), guanine (G), or thymine (T). However, the conformation of the bulge base in the helix (extrahelical or intrahelical) cannot be predicted in the present study.
The thermal denaturation of the duplexes was monitored through the loss of secondary structure via circular dichroism spectroscopy (CD). Thermal melt profiles were conducted in dilute and crowded conditions in order to compare the effect of the cosolutes on the duplex stability. The effects of crowding on the kinetics of hybridization were studied via FAM quenching by fluorescence resonance energy transfer (FRET) (18-20). The strands were labeled with a FRET pair, donor and acceptor fluorophores, which were placed on the same end of the duplex (section 2.2.1). The acceptor, hexachloroflourescein (HEX), was positioned at the 3’ end of strand (A) and the donor, 5’6-carboxyflourescein (FAM), was positioned at the 5’ end of the complementary strand (B). The strategic positioning of the FRET pair on the same side of the duplex maximized the energy transfer rate and quenching of FAM intensity (Scheme 3.1). HEX has an excitation maximum at 538 nm and an emission maximum at 555 nm. FAM has an excitation maximum at 494 nm and an emission maximum at 520 nm.
3.2 Materials and Methods

3.2.1 Materials

Chemicals

The salts, NaCl, KCl, Na$_2$HPO$_4$ and NaH$_2$PO$_4$ were purchased from Sigma-Aldrich. The polymers, poly(ethylene glycol) (PEG) of average molecular weight 3400 (3.4K) and Ficoll type 70 were also purchased from Sigma-Aldrich. The polymers were dissolved in phosphate buffered saline (PBS). The macrosolute concentrations are reported as weight per volume (w/v).

Oligonucleotides

The labeled and unlabeled oligonucleotides were obtained from Integrated DNA Technologies Inc (Coralville, IA). The oligonucleotides were dissolved in PBS (pH 7.4, 137 mM NaCl, 2.6 mM KCl, 10 mM Na$_2$HPO$_4$ and 1.8 mM NaH$_2$PO$_4$). The concentrations of the oligonucleotides were determined by ultra-violet (UV) absorbance on a Beckman spectrophotometer (model DU-640), using the extinction coefficients at 260 nm (Table 3.1). Complementary oligonucleotides were mixed in a 1:1 concentration ratio for the formation of duplex oligonucleotides. The oligonucleotides were annealed by heating at 95 °C for four minutes after which they were cooled to room temperature over a period of two hours.
Table 3.1 – The extinction coefficients (260 nm) for the 5382 sequences. The extinction coefficients were obtained from IDT Inc. and were based on the salt concentrations of PBS buffer.

<table>
<thead>
<tr>
<th></th>
<th>$\varepsilon_{260\text{nm}}^{\text{Unlabeled}}$ (M$^{-1}$ cm$^{-1}$)</th>
<th>$\varepsilon_{260\text{nm}}^{\text{Labeled}}$ (M$^{-1}$ cm$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>WT (A)</td>
<td>170,100</td>
<td>201,700</td>
</tr>
<tr>
<td>WT (B)</td>
<td>140,300</td>
<td>161,300</td>
</tr>
<tr>
<td>A (A)</td>
<td>182,100</td>
<td></td>
</tr>
<tr>
<td>C (A)</td>
<td>176,300</td>
<td>207,900</td>
</tr>
<tr>
<td>C (A) HEX</td>
<td>207,900</td>
<td></td>
</tr>
<tr>
<td>G (A)</td>
<td>180,800</td>
<td></td>
</tr>
<tr>
<td>T (A)</td>
<td>178,200</td>
<td></td>
</tr>
</tbody>
</table>

3.2.2 Measurements

Thermal Denaturation of the 5382 Duplexes

Melt profiles were obtained in dilute and crowded environments. The dilute environment was created with a simple buffer (PBS). The crowded environment was created with the use of a macrosolute (PEG 3.4K or Ficoll 70) at concentrations of 20%. The thermal denaturation of the annealed duplexes was monitored via CD as the change in ellipticity at 245 nm ($\Delta\varepsilon_{245}$) on a Jasco J-710 spectropolarimeter. An external computer-controlled bath was used to control the temperature of the cell holder. All melts were conducted in a 1.0 cm path length quartz cuvette. The samples were overlaid with mineral oil to prevent evaporation during the progress of the experiment. Prior to starting the melt, the samples were equilibrated for 15 minutes at the start temperature. Wavelength scans were collected from 200-300 nm at the beginning and
end of each melt. The melt profiles were obtained by collecting the ellipticity at 245 nm as the temperature increased at a rate of 20 °C per hour. All scans were background corrected against buffer (PBS) or crowded solution. The concentrations used were 4 μM per strand for all duplexes. The values reported are based on the average of three or more individually fit trials. The thermal profiles were created by graphing the change in ellipticity at 245 nm (Δε245) versus temperature.

**Kinetics of Hybridization of 5382-WT and 5382-C**

Hybridization profiles were measured in a dilute environment created with a simple buffer (PBS) and in crowded environments created by the presence of a macrosolute (20% PEG 3.4K or Ficoll type 70). The progress of the hybridization reactions was monitored by FRET, as the quenching of FAM intensity, on a Photon Technologies Incorporated fluorimeter (model QM-2001-4). All kinetics reactions were conducted in a 1.0 cm path length quartz cuvette. A Peltier system was used to maintain the temperature of the cuvette holder at 25 °C. Complementary oligonucleotides were allowed to individually equilibrate at 25 °C for 10 minutes. Upon completion of the equilibration period, the samples were mixed and data collection was initiated. A time-based emission scan was collected at 520 nm (FAM emission peak), with excitation at 494 nm at 25 °C. An automatic shutter was employed to prevent photo-bleaching of the dyes. Scans were collected by opening the shutter for five seconds and collecting data at a rate of 0.2 points per second. Following each five
second collection window was a 15 second period during which no data was collected and the shutter was closed. The total collection times, 30 or 60 minutes, were dependent on the sequences and conditions. Excitation and emission slit widths were 3 and 6 nm respectively. A 480-499 nm band-pass filter obtained from Chroma Technology Corporation (Rockingham, VT) was used in the excitation beam to reduce light scattering. Excitation correction was used to correct for fluctuations in the lamp intensity. All scans were background corrected against buffer (PBS) or crowded solution. The kinetics of hybridization profiles were created by graphing the normalized intensity at 520 nm (I₅₂₀) versus time. The concentration used was 5 nM per strand for all duplexes. The values reported are based on the fit of an average data set which was obtained from three or more trials.

3.2.3 Data Analysis

Analysis of the 5382 Melt Profiles

The melting profiles of the 5382 duplexes were obtained by plotting the change in ellipticity at 245 nm (Δɛ₂₄₅) versus temperature in Kelvin (K). The melt profiles were fit with an expression for two-state melting of a non-self complementary sequence with Kaleidagraph 3.6. The formation of the 5382 duplexes can be expressed as (21,22):

\[
A + B \rightleftharpoons AB . \quad \text{(Equation 3.1)}
\]
where two strands comprise the final structure. Hence, the molecularity (n) of the 5382 sequences is two. The equilibrium constant (\(K_{eq}(T)\)) for the association described above can be expressed as:

\[
K_{eq}(T) = \frac{[AB]}{[A][B]} = \frac{\alpha}{\left(\frac{C_T}{n}\right)^{n-1} (1-\alpha)^n} = \frac{2\alpha}{(C_T(1-\alpha)^2} \quad \text{(Equation 3.2)}
\]

where \(C_T\) represents the total strand concentration and \(\alpha\) is the fraction of double stranded DNA (21,22). Mathematica 6.0 was used to solve Equation 3.2 in terms of \(\alpha\). The negative form of the root was real in the range of the experimental data.

\[
\alpha(T) = \frac{1+C_TK_{eq}(T)-\sqrt{1+2C_TK_{eq}(T)}}{C_TK_{eq}(T)} \quad \text{(Equation 3.3)}
\]

The van’t Hoff expression for \(K_{eq}(T)\) is defined as:

\[
K_{eq}(T) = \exp\left(-\frac{\Delta G}{RT}\right) = \exp\left(-\frac{\Delta H}{RT} + \frac{\Delta S}{R}\right) \quad \text{(Equation 3.4)}
\]

where \(\Delta G\) is the change in free energy, \(\Delta H\) is the change in enthalpy, \(\Delta S\) refers to the change in entropy and \(R\) is the gas constant, 1.987 cal mol\(^{-1}\) K\(^{-1}\) (21). Substitution of the van’t Hoff expression for \(K_{eq}\) into Equation 3.3 yields an expression for \(\alpha\) in terms of \(\Delta H\) and \(\Delta S\).

\[
\alpha(T) = \frac{1+C_T\exp\left(-\frac{\Delta H}{RT} + \frac{\Delta S}{R}\right)}{C_T\exp\left(-\frac{\Delta H}{RT} + \frac{\Delta S}{R}\right)} \quad \text{(Equation 3.5)}
\]
The change in ellipticity at 245 nm at a given temperature ($\Delta \varepsilon_{245}(T)$) can be defined by the following expression:

$$\Delta \varepsilon_{245}(T) = \alpha [\varepsilon_{245}(T_{\text{min}}) + m_{\text{ds}}(T - T_{\text{min}})] + (1 - \alpha)[\varepsilon_{245}(T_{\text{max}}) - m_{\text{ss}}(T_{\text{max}} - T)]$$

(Equation 3.6)

where $\varepsilon_{245}(T_{\text{min}})$ is the average measured ellipticity at the temperature minimum ($T_{\text{min}}$) and $\varepsilon_{245}(T_{\text{max}})$ is the average measured ellipticity at the temperature maximum ($T_{\text{max}}$) (22). The $m_{\text{ds}}$ and $m_{\text{ss}}$ are the average slopes for the first and last 10-15 °C temperature points respectively. The $m_{\text{ds}}$ represents the temperature dependence of the ellipticity of the double-stranded (ds) DNA and $m_{\text{ss}}$ represents the temperature dependence of ellipticity the single-stranded (ss) DNA (22). This procedure corrected for the temperature dependence of the upper and lower baselines.

The $\alpha$ term defined in Equation 3.5 was substituted into Equation 3.6 and used to analyze the melt profiles for the 5382 duplexes (equation not shown). The $\Delta H$ and $\Delta S$ were obtained from the curve fit analysis and were subsequently used to calculate the melting point ($T_m$). The following general equation for $T_m$ can be applied to any non-self-complementary system of any molecularity (21).

$$\frac{1}{T_m} = \frac{(n-1)R}{\Delta H} \ln C_T + \frac{[\Delta S - (n-1)R\ln(2n)]}{\Delta H}$$

(Equation 3.7)

The $T_m$ expression for the 5382 (n=2) system is:
\[ T_m = \frac{\Delta H}{\Delta S + R \ln \left( \frac{C^T}{4} \right)} \]. \hspace{1cm} \text{(Equation 3.8)}

The Gibbs free energy (\(\Delta G\)) was calculated at 37 \(^\circ\)C because it is physiologically relevant (22).

\[ \Delta G = \Delta H - T \Delta S = -RT \ln K_{eq} \hspace{1cm} \text{(Equation 3.9)} \]

The melt profiles were individually fit to obtain \(\Delta H\), \(\Delta S\) and calculate \(\Delta G_{37^\circ C}\) and \(T_m\) values. The reported \(\Delta H\), \(\Delta S\), \(\Delta G_{37^\circ C}\) and \(T_m\) are average values based on a minimum of 3 trials. The error for parameters was calculated as standard error of the mean.

**Analysis of the 5382 Hybridization Profiles**

The kinetics of duplex formation is a second order reaction where \(A + B \rightarrow AB\) and the rate can be described as:

\[ r = k[A][B] \]. \hspace{1cm} \text{(Equation 3.10)}

Although \(A\) and \(B\) are two different reactants, since the single strands were mixed in equimolar ratios, Equation 3.10 can be simplified using \([S]\) which represented the single strands that comprise the duplex (23).

\[ r = [S]^2 \hspace{1cm} \text{(Equation 3.11)} \]

Integration of the rate law and separation of variables yields an expression which describes \([S]_t\) at any time (t) where at \(t=0\), \([S]_t\) is equal to the sum of \([A]_0\) and \([B]_0\) (23).
\[ [S]_t = \frac{[S]_0}{1 + k_t[S]_0}. \]  \hspace{1cm} \text{(Equation 3.12)}

The \([S]_t\) is then converted to \(I_{520}(t)\), the normalized intensity at 520 nm at time \(t\).

\[ I_{520}(t) = f_1 \left( \frac{[S]_0}{1 + k_t[S]_0} \right) + f_2 \left( [S]_{\text{total}} - \frac{[S]_0}{1 + k_t[S]_0} \right) \]  \hspace{1cm} \text{(Equation 3.13)}

In Equation 3.13, \(f_1\) is the signal contribution from the single stranded DNA and \(f_2\) is the signal contribution from the double-stranded DNA, \(t\) is time in seconds, and \(k\) is the rate constant in M\(^{-1}\)s\(^{-1}\). The hybridization profiles were fit with Kaleidagraph 3.6.

Prior to the curve fitting, the data was normalized using the following equation:

\[ I_{520}(t) = \frac{I_{520}(t) - I_{520}(\text{min})}{I_{520}(\text{max}) - I_{520}(\text{min})} \]  \hspace{1cm} \text{(Equation 3.14)}

where \(I_{520}(\text{min})\) and \(I_{520}(\text{max})\) refer to the minimum and maximum intensity at 520 nm respectively. Three or more normalized data sets were averaged to obtain the rate constants of hybridization.

3.3 Results

3.3.1 Macrosolute Effects on the Thermodynamics of the 5382 Duplexes

Melt profiles of all duplexes were collected in dilute environment which was created with a simple buffer (PBS) and in crowded environments created by the presence of a macrosolute (20% PEG 3.4K or Ficoll 70). Macrosolute selection was based on previous effects noted with other systems (chapter 2). The thermodynamic stabilities of the 5382-WT and bulge duplexes were studied by heating the duplexes and monitoring the loss of secondary structure via CD at 245 nm. Figure 3.1 shows
representative wavelength scans collected at the start and end of melts for the 5382-WT duplex in crowded and dilute environments. The wavelength scans show the characteristic peak minimum and maximum at 245 and 280 nm respectively, which indicate that the 5382-WT duplex is in B-form (24). The monitoring wavelength of 245 nm, was selected based on the large change in signal at that wavelength upon total denaturation of the duplexes.

The temperature at which half of the duplex is lost is the melt temperature ($T_m$). In Figure 3.1 the two-state melting profiles for the 5382-WT duplex are shown in dilute and crowded environments. In the dilute environment, the 5382-WT duplex had a $T_m$ of 67.3 °C (Table 3.2). In the presence of 20% PEG 3.4K the $T_m$ increased to 69.3 °C. The difference in the $T_m$ of 5382-WT in the crowded compared to the dilute environment, $\Delta T_m$, is a measure of the functional significance of the macrosolutes. As shown in Table 3.3, the greater $\Delta T_m$ for the 5382-WT was 2.0 °C in the presence of 20% PEG 3.4K. In the presence of 20% Ficoll, minimal stabilization was noted for the 5382-WT duplex. The Gibbs free energy measures the stability of a structure and is a thermodynamic parameter that can be utilized as a standard for comparing different structures. The difference in the free energy ($\Delta \Delta G_{37C}$), is a measure of the macrosolute induced stabilization on a duplex. The $\Delta \Delta G_{37C}$ for the 5382-WT duplex was -0.9 kcal mol$^{-1}$ in the presence of 20% PEG 3.4K.

In PBS, the $T_m$ of all the bulge duplexes was lower than that of the 5382-WT duplex which demonstrated the disruption to the helix that results from the single base
bulge. However, the bulge duplexes maintained a B-form structure as indicated by the wavelength scans which contain the characteristic minimum and maximum peaks (figures 3.2 to 3.5) (24). The bulge duplexes can be organized according the destabilizing of the bulge helix relative to the WT. The most destabilized duplex was 5382-C, the 5382-G and 5382-T duplexes were moderately destabilized and finally the 5382-A duplex was the least destabilized as compared to the parent duplex in PBS. All of the bulge duplexes were stabilized in the presence of the cosolutes however to varying extents. Similar to the 5382-WT duplex, all of the bulge duplexes exhibited marginal stabilization in the presence of 20% Ficoll.

The incorporation of a cytosine was the most disruptive bulge base to the duplex. The $T_m$ of 5382-C in PBS was 54.9 °C, a 12.4 °C decrease in the $T_m$ as compared to the parent duplex. Although, 5382-C had the lowest $T_m$ of all the bulge duplexes, it exhibited the greatest crowding effects with a $\Delta T_m$ of 4.7 °C in the presence of 20% PEG 3.4K. Additionally, the $\Delta \Delta G_{37^\circ C}$ for the 5382-C duplex was -2.0 kcal mol$^{-1}$ in the presence of 20% PEG 3.4K. The macrosolute induced stabilization of the 5382-C duplex was twice that of the 5382-WT duplex in 20% PEG 3.4K.

The incorporation of a guanine or thymine base resulted in median disruption to the duplex as indicated by the comparison of the $T_m$’s of the bulge and parent duplexes in PBS. The respective $T_m$’s of the 5382-G and 5382-T duplexes in PBS was 55.2 and 57.0 °C which are 10.3 and 12.1 °C decreases in $T_m$ as compared to the 5382-WT parent duplex (Table 3.2). The crowding effects that 5382-G and 5382-T duplexes exhibited
were also in the median range (figures 3.3 and 3.4). In the presence of 20% PEG 3.4K the 5382-G and 5382-T duplexes had similar ΔTm’s of 2.9 and 3.2 °C respectively (Table 3.3). The ΔΔG_{37°C} for the 5382-G duplex was -1.3 kcal mol⁻¹, and -1.2 kcal mol⁻¹ for the 5382-T duplex, which demonstrates a greater stabilization of these bulge duplexes as compared to the 5382-WT.

The incorporation of the adenine base resulted in the least disruption to the helix (Figure 3.5). The Tm of the duplex was 57.8 °C, a 9.5 °C decrease in Tm as compared to the parent duplex (Table 3.2). In the presence of PEG 3.4K, the ΔTm was 2.5 °C, a value similar to the stabilization exhibited by the 5382-WT duplex under the same conditions (Table 3.3). The ΔΔG_{37°C} was -1.0 kcal mol⁻¹ in the presence of PEG 3.4K. Hence, the overall stabilization of the 5382-A duplex was equivalent to the stabilization noted for the parent duplex.
Figure 3.1 – Representative wavelength scans and melt profiles of the 5382-WT duplexes in dilute and crowded environments. The blue and red circles represent wavelength scans performed in PBS at 25 and 95 °C respectively. The melt profiles in PBS, 20% Ficoll and PEG 3.4K were fit as shown by the black, orange and blue lines respectively. (A) PBS scans (B) PBS melts (C) 20% Ficoll 70 scans (D) 20% Ficoll 70 melts (E) 20% PEG 3.4K scans and (F) 20% PEG 3.4K melts.
Figure 3.2 – Representative wavelength scans and melt profiles of the 5382-C duplexes in dilute and crowded environments. The blue and red circles represent wavelength scans performed in PBS at 25 and 95 °C respectively. The melt profiles in PBS, 20% Ficoll and PEG 3.4K were fit as shown by the black, orange and blue lines respectively. (A) PBS scans (B) PBS melts (C) 20% Ficoll 70 scans (D) 20% Ficoll 70 melts (E) 20% PEG 3.4K scans and (F) 20% PEG 3.4K melts.
Figure 3.3 – Representative wavelength scans and melt profiles of the 5382-T duplexes in dilute and crowded environments. The blue and red circles represent wavelength scans performed in PBS at 25 and 95 °C respectively. The melt profiles in PBS, 20% Ficoll and PEG 3.4K were fit as shown by the black, orange and blue lines respectively. (A) PBS scans (B) PBS melts (C) 20% Ficoll 70 scans (D) 20% Ficoll 70 melts (E) 20% PEG 3.4K scans and (F) 20% PEG 3.4K melts.
Figure 3.4 – Representative wavelength scans and melt profiles of the 5382-G duplexes in dilute and crowded environments. The blue and red circles represent wavelength scans performed in PBS at 25 and 95 °C respectively. The melt profiles in PBS, 20% Ficoll and PEG 3.4K were fit as shown by the black, orange and blue lines respectively. (A) PBS scans (B) PBS melts (C) 20% Ficoll 70 scans (D) 20% Ficoll 70 melts (E) 20% PEG 3.4K scans and (F) 20% PEG 3.4K melts.
Figure 3.5 – Representative wavelength scans and melt profiles of the 5382-A duplexes in dilute and crowded environments. The blue and red circles represent wavelength scans performed in PBS at 25 and 95 °C respectively. The melt profiles in PBS, 20% Ficoll and PEG 3.4K were fit as shown by the black, orange and blue lines respectively. (A) PBS scans (B) PBS melts (C) 20% Ficoll 70 scans (D) 20% Ficoll 70 melts (E) 20% PEG 3.4K scans and (F) 20% PEG 3.4K melts.
<table>
<thead>
<tr>
<th>5382-WT</th>
<th>PBS</th>
<th>20% Ficoll 70</th>
<th>20% PEG 3.4K</th>
</tr>
</thead>
<tbody>
<tr>
<td>ΔH (kcal mol⁻¹)</td>
<td>-135.4 ± 0.3</td>
<td>-136.9 ± 0.5</td>
<td>-137.3 ± 0.5</td>
</tr>
<tr>
<td>ΔS (cal mol⁻¹ K⁻¹)</td>
<td>-372.5 ± 0.9</td>
<td>-375.3 ± 1.9</td>
<td>-374.8 ± 1.8</td>
</tr>
<tr>
<td>T_m (°C)</td>
<td>67.3 ± 0.2</td>
<td>67.9 ± 0.5</td>
<td>69.3 ± 0.3</td>
</tr>
<tr>
<td>ΔG₃⁷°C (kcal mol⁻¹)</td>
<td>-20.2 ± 0.1</td>
<td>-20.5 ± 0.1</td>
<td>-21.1 ± 0.1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>5382-C</th>
<th>PBS</th>
<th>20% Ficoll 70</th>
<th>20% PEG 3.4K</th>
</tr>
</thead>
<tbody>
<tr>
<td>ΔH (kcal mol⁻¹)</td>
<td>-120.5 ± 0.2</td>
<td>-121.3 ± 2.0</td>
<td>-127.6 ± 1.8</td>
</tr>
<tr>
<td>ΔS (cal mol⁻¹ K⁻¹)</td>
<td>-341.0 ± 0.9</td>
<td>-341.7 ± 5.9</td>
<td>-354.4 ± 5.1</td>
</tr>
<tr>
<td>T_m (°C)</td>
<td>54.9 ± 0.3</td>
<td>56.5 ± 0.1</td>
<td>59.6 ± 0.3</td>
</tr>
<tr>
<td>ΔG₃⁷°C (kcal mol⁻¹)</td>
<td>-14.7 ± 0.1</td>
<td>-15.3 ± 0.1</td>
<td>-16.7 ± 0.2</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>5382-T</th>
<th>PBS</th>
<th>20% Ficoll 70</th>
<th>20% PEG 3.4K</th>
</tr>
</thead>
<tbody>
<tr>
<td>ΔH (kcal mol⁻¹)</td>
<td>-124.5 ± 0.5</td>
<td>-125.0 ± 1.4</td>
<td>-126.0 ± 0.4</td>
</tr>
<tr>
<td>ΔS (cal mol⁻¹ K⁻¹)</td>
<td>-350.9 ± 2.0</td>
<td>-352.2 ± 4.6</td>
<td>-351.8 ± 0.5</td>
</tr>
<tr>
<td>T_m (°C)</td>
<td>57.0 ± 0.3</td>
<td>57.3 ± 0.3</td>
<td>60.2 ± 0.5</td>
</tr>
<tr>
<td>ΔG₃⁷°C (kcal mol⁻¹)</td>
<td>-15.6 ± 0.1</td>
<td>-15.7 ± 0.1</td>
<td>-16.8 ± 0.2</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>5382-G</th>
<th>PBS</th>
<th>20% Ficoll 70</th>
<th>20% PEG 3.4K</th>
</tr>
</thead>
<tbody>
<tr>
<td>ΔH (kcal mol⁻¹)</td>
<td>-122.1 ± 0.5</td>
<td>-125.1 ± 0.6</td>
<td>-125.0 ± 0.1</td>
</tr>
<tr>
<td>ΔS (cal mol⁻¹ K⁻¹)</td>
<td>-345.9 ± 1.2</td>
<td>-353.4 ± 2.0</td>
<td>-351.3 ± 0.6</td>
</tr>
<tr>
<td>T_m (°C)</td>
<td>55.2 ± 0.4</td>
<td>56.5 ± 0.3</td>
<td>58.1 ± 0.3</td>
</tr>
<tr>
<td>ΔG₃⁷°C (kcal mol⁻¹)</td>
<td>-14.8 ± 0.2</td>
<td>-15.5 ± 0.1</td>
<td>-16.1 ± 0.1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>5382-A</th>
<th>PBS</th>
<th>20% Ficoll 70</th>
<th>20% PEG 3.4K</th>
</tr>
</thead>
<tbody>
<tr>
<td>ΔH (kcal mol⁻¹)</td>
<td>-124.0 ± 0.3</td>
<td>-125.1 ± 0.5</td>
<td>-125.9 ± 0.4</td>
</tr>
<tr>
<td>ΔS (cal mol⁻¹ K⁻¹)</td>
<td>-348.7 ± 1.1</td>
<td>-352.0 ± 1.5</td>
<td>-351.4 ± 1.4</td>
</tr>
<tr>
<td>T_m (°C)</td>
<td>57.8 ± 0.5</td>
<td>57.8 ± 0.1</td>
<td>60.3 ± 0.1</td>
</tr>
<tr>
<td>ΔG₃⁷°C (kcal mol⁻¹)</td>
<td>-15.9 ± 0.2</td>
<td>-16.0 ± 0.1</td>
<td>-16.9 ± 0.1</td>
</tr>
</tbody>
</table>

Table 3.2 - The thermodynamic parameters of hybridization for the 5382 duplexes in PBS and macrosolutes. All values shown are based on the average of at least three trials. Reported errors were calculated as the standard error of the mean.
Table 3.3 – The changes in $T_m$ and $\Delta \Delta G_{37C}$ for the 5382 duplexes in the presence of macrosolutes. These values were calculated as the difference of the parameter values in crowded versus PBS (dilute). Reported error values were tabulated using propagation of error.

<table>
<thead>
<tr>
<th></th>
<th>20% Ficoll 70 - PBS</th>
<th>20% PEG 3.4K - PBS</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>5382-WT</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\Delta T_m$</td>
<td>0.6 ± 0.5</td>
<td>2.0 ± 0.4</td>
</tr>
<tr>
<td>$\Delta \Delta G_{37C}$</td>
<td>-0.3 ± 0.1</td>
<td>-0.9 ± 0.1</td>
</tr>
<tr>
<td><strong>5382-C</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\Delta T_m$</td>
<td>1.6 ± 0.3</td>
<td>4.7 ± 0.5</td>
</tr>
<tr>
<td>$\Delta \Delta G_{37C}$</td>
<td>-0.6 ± 0.1</td>
<td>-2.0 ± 0.2</td>
</tr>
<tr>
<td><strong>5382-T</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\Delta T_m$</td>
<td>0.3 ± 0.4</td>
<td>3.2 ± 0.6</td>
</tr>
<tr>
<td>$\Delta \Delta G_{37C}$</td>
<td>-0.1 ± 0.1</td>
<td>-1.2 ± 0.2</td>
</tr>
<tr>
<td><strong>5382-G</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\Delta T_m$</td>
<td>1.3 ± 0.5</td>
<td>2.9 ± 0.5</td>
</tr>
<tr>
<td>$\Delta \Delta G_{37C}$</td>
<td>-0.7 ± 0.2</td>
<td>-1.3 ± 0.2</td>
</tr>
<tr>
<td><strong>5382-A</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\Delta T_m$</td>
<td>0.0 ± 0.5</td>
<td>2.5 ± 0.5</td>
</tr>
<tr>
<td>$\Delta \Delta G_{37C}$</td>
<td>-0.1 ± 0.2</td>
<td>-1.0 ± 0.2</td>
</tr>
</tbody>
</table>

3.3.2 Macrosolute Effects on the Kinetics of Hybridization of the 5382-WT and 5382-C Duplexes

The presence of macrosolutes can also have an effect on the kinetics of duplex formation. A consequence of the macrosolute presence is an increase in the forward rate of reaction (25). Although the macrosolutes increase the viscosity of the system, an increase in the chemical activity opposes the viscosity effects (10,25,26). Previously we studied the kinetics of formation of duplexes with single base mismatches (chapter 2). The crowding induced enhancement of the rate constant was observed for both
mismatched and WT duplexes. Here we present the effects of macrosolute on the rate constant of formation of a duplex containing a single base bulge. The kinetics of 5382 duplex formation were studied under dilute conditions (PBS) and crowded conditions created by the presence of a macrosolute (20% PEG 3.4K or Ficoll 70). Coseolute selection and concentration was based on the results observed in the thermodynamic studies of the 5382 duplexes (section 3.3.1).

The duplexes for the kinetic studies were also selected based on previous thermodynamic results (section 3.3.1). In the presence of 20% PEG 3.4K, the 5382-C duplex was differentially stabilized over the 5382-WT duplex. The kinetic studies aimed at investigating whether the kinetics of duplex formation for 5382-C could be enhanced to a greater extent than the 5382-WT duplex. The strands of the duplex were labeled with a HEX and FAM fluorophores in order to monitor the kinetics of duplex formation via the quenching of FAM by FRET (section 2.2). As time and duplex formation progressed, the strands moved closer together on average and the intensity of the FAM peak decreased as a result of the fluorophores proximity (figures 3.6). The intensity of the donor fluorophore (FAM) was monitored against time. The data was fit to a second order rate equation to obtain the rate constants for the reactions.

The rate constants for the duplex formation in the presence and absence of macrosolutes are shown in Table 3.4. Both of the duplexes have rate constants that are in agreement with other reported values of duplex formation (27). In PBS, the rate constant of duplex formation for the 5382-WT was $9.7 \times 10^5 \text{ M}^{-1}\text{s}^{-1}$. The 5382-C
duplex had a rate constant for hybridization of $4.1 \times 10^5 \text{M}^{-1}\text{s}^{-1}$, over 2-fold slower than the 5382-WT duplex. Neither the 5382-WT nor 5382-C strands showed any significant increase in hybridization rates in the presence of Ficoll 70. However, in the presence of 20% PEG 3.4K the 5382-WT and 5382-C hybridization rates were increased by 1.8 and 3-fold respectively.

![Figure 3.6](image)

Figure 3.6 - The kinetics of formation for the 5382-WT and 5382-C duplexes in dilute and crowded environments. (A) 5382-WT and (B) 5382-C hybridization profiles in PBS (black), 20% Ficoll 70 (green) and 20% PEG 3.4K (blue). The data points and error bars shown are based on the average of at least three trials. The solid lines represent a second order curve-fit through the data.

<table>
<thead>
<tr>
<th></th>
<th>5382-WT $K$ ($10^5 \text{M}^{-1}\text{s}^{-1}$)</th>
<th>5382-C $K$ ($10^5 \text{M}^{-1}\text{s}^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1X PBS</td>
<td>9.7 ± 0.1</td>
<td>4.1 ± 0.1</td>
</tr>
<tr>
<td>20% Ficoll 70</td>
<td>10.6 ± 0.1</td>
<td>5.4 ± 0.1</td>
</tr>
<tr>
<td>20% PEG 3.4K</td>
<td>18.0 ± 0.2</td>
<td>12.4 ± 0.1</td>
</tr>
</tbody>
</table>

Table 3.4 The rate constants of formation for the 5382-WT and 5382-C duplexes in the presence of macrosolutes. Values are based on the average of three or more trials. Reported error values are the standard error of predication obtained from the fit.
3.4 Discussion

3.4.1 Crowding Mediated Stabilization of the Bulge Duplexes

Effects of a single base bulge on a duplex

The system studied here consisted of a 17 base pair parent duplex and four duplexes with single base bulges. The presence of the bulge bases resulted in disruption of all the bulge duplexes as compared to the parent duplex, 5382-WT. However, the disruption varied as noted in the differences in $T_m$ of the 5382-WT duplex versus the bulge duplexes in PBS. In PBS, the bulge duplexes had $T_m$’s which were 9.5 to 12.4 °C lower than that of 5382-WT. The bulge base duplexes, arranged in order of least to most destabilized (relative to 5382-WT) are 5832-A, 5382-T, 5382-G and 5382-C. Hence, the data presented here is in agreement with the findings of Ke and Wartell who noted that pyrimidines are the most disruptive bulges to a duplex and purines are the least disruptive (28).

In a subsequent publication, Ke and Wartell categorized the bulge bases into two groups (7). The first group of bulge bases is where the inserted base is different than its neighboring bases. The second group contains bulge bases that are identical to at least one of its neighboring bases. The bulges in group II are not as disruptive to the helix as the those in group I. The work of Woodston and Crothers reaffirms the findings put forth by Wartell and Ke (29). Classification of the 5382 duplexes according to the guidelines described by Wartell and Ke places the 5382-G and 5382-C duplexes into group I. The 5382-G and 5382-C duplexes had $T_m$’s that were 12.1 and
12.4 °C lower than the 5382-WT duplex respectively. The 5382-A and 5382-T bulge duplexes are in group II. The 5382-A and 5382-T duplexes were the least destabilized with a 9.5 and 10.3 °C decrease in $T_m$ relative to the 5382-WT duplex respectively. The group II bases were not as disruptive as the group I bases. Hence, our results are in agreement with the work of Ke and Wartell.

Our findings are further confirmed by the work of Santalucia and Hicks who noted that C and A bulges are the most and least destabilizing to a duplex respectively (30). Santalucia and Hicks also noted that G and T bulges were intermediate in their destabilization effect. The differences in bulge base consequences on a helix are attributed to positional degeneracy of group I bulges (7). It is also postulated that aside from the group I versus group II classification, the conformational freedom of the purine versus pyrimidine rings is also a factor in how they affect a helix structure (31).

**Differential Stabilization of the 5382-C Duplex in a Crowded Environment**

Despite the disruption that the bulge duplexes exhibited, all of the duplexes were stabilized against thermal denaturation in the presence of the macrosolutes (20% PEG 3.4K and Ficoll 70). The $\Delta H$ and $\Delta S$ values obtained were comparable to those of other 17-mer sequences with single base bulges (4,5,17). Our results are also in agreement with Ochuci and colleagues who found a 4 kcal mol$^{-1}$ difference in $\Delta G_{37C}$ between the parent duplex and a single bulge base duplex (17). The average difference
in $\Delta G_{37^\circ C}$ between the parent duplex and the 5382 bulge duplexes in PBS was 5 kcal mol$^{-1}$.

The effects of crowding are dependent on both the size and shape of the cosolutes (32). Of the macrosolutes studied here, the Ficoll exhibited little effect while the PEG 3.4K exhibited significant stabilization of the duplexes. The smaller Ficoll effects may be explained by the fact that it is a rod-like shaped cosolute. It has been shown that rod-like cosolutes have less effect than globular like cosolutes, such as PEG 3.4K (33). Here we report a 4.6 $^\circ C$ increase in $T_m$ for the 5382-C duplex in the presence of 20% PEG 3.4K. The effects of PEG 3.4K on the thermal stability of the 5382 duplexes is in agreement with expectations based on previously reported values (33,34). Abraham Minsky studied the effects of macrosolutes on 20 base pair duplexes and found an average 3-5 $^\circ C$ increase in the $T_m$ of duplexes (33,34). Chaires and Spink also studied duplexes in the presence of macrosolutes and reported $T_m$ increases of 3-5 $^\circ C$ on average (35,36).

The $\Delta \Delta G_{37^\circ C}$ of the 5382-WT duplex was -0.9 kcal mol$^{-1}$ in the presence of PEG 3.4K. The $\Delta \Delta G_{37^\circ C}$ of the 5382-C duplex was -2.0 kcal mol$^{-1}$ under the same conditions. The difference between the $\Delta \Delta G_{37^\circ C}$ of the two duplexes is -1.1 kcal mol$^{-1}$ and demonstrates the differential stabilization of the 5382-C duplex over its parent duplex. The 5382-C duplex was also stabilized to a greater extent than any of the other bulge duplexes. Additionally, all of the bulge duplexes were stabilized to a greater extent than the WT duplex. The 5382-C duplex contained the bulge base that was the
most disruptive to the helix. The results presented here contrast with those of the 
BRCA1-185 thermodynamic studies. In the BRCA1-185 system of duplexes, a CC and 
GG base pair were selected for incorporation into a duplex as they are the most and 
least disruptive to a helix (37). In the presence of macrosolutes, differential 
stabilization was observed for the WT duplex over both single mismatched duplexes 
(chapter 2).

The differential stabilization of 5382-C may be due to the conformation of the 
bulge base within the helix. However, it is not possible to determine whether the 5382 
duplexes contain intra- or extrahelical bases within the scope of this research project. 
However, some general assumptions can be drawn from literature in order to interpret 
the effects observed. Marky and colleagues studied the change in volume for single 
bulge A, T and WT duplexes (4). Upon duplex formation, water is taken up and 
results in an increase in the change in volume (38). Marky found that the bulge 
duplexes resulted in more hydrated structures with the bulge containing duplexes 
exhibiting changes in volume of up to 1.5 greater than the parent duplex. The 
increased hydration is believed to aid in stabilizing the bulge duplex (4).

The duplexes that Marky studied contained two single base bulges both of which 
were extrahelical. Intrahelical bases would have resulted in smaller changes in volume 
(4). It is possible that the 5382-C duplex is extrahelical and the structural consequences 
of the bulge may have led to differential crowding effects. We speculate that the other 
bulge duplexes contained bases which were all intrahelical and hence they had similar
effects as the WT duplex. Furthermore, if 5382-C were extrahelical, differences in hydration between it and the other duplexes might also help explain the differential crowding effects. Further studies, such as nuclear magnetic resonance (NMR) would be required in order to determine the extrahelical or intrahelical nature of the bulge bases.

### 3.4.2 Macrosolute Enhancement of the Rate Constant of Hybridization

In the kinetic assays, the cosolutes were added to study the effects of crowding on duplex formation. Strand, cosolute, and duplex selection were based on the effects noted in the thermodynamic studies of 5382 duplexes. The 5382-WT and 5382-C were selected for the kinetic studies since 5382-C was differentially stabilized over the 5382-WT duplex. The rate constants of duplex formation presented here are in agreement with the typical rate constants of duplex formation for DNA, about $10^5$ to $10^7$ M$^{-1}$s$^{-1}$ (27,39). The rate limiting step for duplex formation is the nucleation site, which is the base pairing of the 2-3 bases (25). The energy barrier for duplex formation is determined by the rate limiting step. This view of duplex formation is commonly called the “zip-up” or “zipper model” (40). In PBS, the rate constant of 5382-WT duplex formation was approximately 2-fold faster than that of 5382-C. This two-fold difference in the rate constant can be accounted for by the zipper model of hybridization. It has been postulated that hybridization of single bulge containing duplexes occurs upon the formation of a nucleation site of up to 4-5 bases (39). The
increased number of base pairs in the nucleation site could result in a slower rate of duplex formation for bulge base containing duplexes as shown here.

In the presence of macrosolutes, the forward rate of duplex formation for the 5382-WT and 5382-C duplexes was increased in PEG 3.4K. The 5382-C and the 5382-WT duplexes exhibited marginal kinetic enhancement in the presence of 20% Ficoll. Similarly, in the thermodynamic studies, Ficoll effects were not as significant as those of PEG 3.4K. The rate constant of the 5382-WT formation in PEG 3.4K versus PBS was 1.8-fold larger than in PBS. The rate constant of the 5382-C formation in PEG 3.4K was 3-fold faster. Although a slightly larger crowding effect was noted for the 5382-C duplex versus the 5382-WT duplex, the difference is not substantial enough to conclude that any significant difference in the mechanism of formation occurs.

3.5 Conclusions

Single base bulges are a common, and naturally occurring structure (5). These structures can arise from recombination in sequences such as homopolymers or from errors during DNA replication (5). Further understanding of how bulge bases are influenced by their naturally crowded environment can aid in deciphering their biological role. Bulged bases are repaired by the methyl-directed mismatch repair pathway (6). Up to three bulge bases can be repaired as efficiently as some
mismatched base pairs (6). Bulge region are also recognition sites for proteins and RNA duplexes or RNA-DNA duplexes (37). However, all of these studies have been conducted in dilute conditions and as demonstrated here macromolecular crowding may result in differential stabilization of one bulge base versus others. In order to further comprehend the biological and structural consequences of bulge bases, knowledge of their structural and chemical properties in crowded environments is crucial.

Here we have demonstrated that a single base bulge destabilized the 5382 bulge duplexes as compared to the parent 5382-WT duplex. However, the crowding-mediated stabilization of all the duplexes was observed in the presence macrosolutes. The macrosolute PEG 3.4K resulted in the greater thermal stability enhancement for all duplexes over the Ficoll 70. The greatest crowding-mediated stabilized bulge duplex was 5382-C, the most destabilized duplex. In the presence of macrosolute, the 5382-C duplex was differentially stabilized over the 5382-WT duplex. The difference in stabilization for 5382-C could possibly be due to the orientation of the bulge within the helix (4). An enhancement of the rate constant of duplex formation was observed for both the 5382-WT and 5382-C duplexes in the presence of 20% PEG 3.4K.
References


Chapter 4 – Macrosolute Effects on Unimolecular Systems

4.1 Introduction

In addition to the common duplex form, DNA can exist in a variety of structures including stem-loop, triplex and quadruplex (1). A stem-loop structure forms from a single strand of DNA which has regions of complementarily which forms a stem (2). The remaining bases, which are not involved in the stem, can be found in the loop of the structure. These unimolecular structures are referred to as molecular beacons when donor and quencher fluorophores are placed on the ends of the stem (3). Molecular beacons are not fluorescent when they are in the stem-loop configuration, because of the proximity of the fluorophores (4). However, upon hybridization to its complementary target, the molecular beacon becomes fluorescent. Although the design of molecular beacons is simple, they are robust probes of hybridization, binding to their target nucleic acids both in vitro and in vivo (5).

Molecular beacons have been used to probe the structure of single and double stranded DNA and to detect DNA ligation in real time (6,7). Additionally, molecular beacons have been utilized to identify single-nucleotide polymorphisms, HIV-1 and other retroviruses (8,9). Specifically, the real-time detection of mRNA expression can be used to quantify the expression levels of cancer-linked genes (10,11). Despite their numerous applications, little is known about the effects of a crowded environment on the intramolecularly stabilized stem-loop structures. The crowding mediated effects of macrosolutes on duplexes, which are intermolecularly stabilized, have been studied (12-
18). For example, it has been shown that a 15-20 base pair duplex can be stabilized 3-5 °C in a crowded environment (18). Therefore, the goal of the stem-loop studies was to determine if the crowding-mediated thermodynamic effects on bimolecular systems are equivalent to effects on unimolecular systems. Additionally, the effects of crowding on stem-loop and target kinetics was also examined.

**Cyclin D1 and Survivin Stem-loop Systems**

The sequences for the unimolecular thermodynamic and duplex kinetic studies, cyclin D1 and survivin, are biologically relevant and heterogeneous. Both sequences have been used in the detection of their respective mRNA targets in living cells as a measure of gene expression levels (10,11). Survivin is a member of the inhibitor of apoptosis (IAP) protein family (19). Although survivin is not detectable in most human adult tissues, it is expressed during fetal development (20). Increased levels of survivin have been linked with several cancers, including 70% of breast cancers (21). Survivin and the other members of the IAP family are believed to be involved in the apoptotic resistance of tumor cells (19). The cyclin D1 gene encodes a protein that is involved in the regulation of the G1 phase of the cell-cycle (22). Like survivin, cyclin D1 is not generally expressed in normal adult tissues (10). The overexpression of cyclin D1 has been noted in 50-80% of breast cancer tissues (10).
In the present study, the effects of crowding on the thermodynamics of stem-loops and their duplexes were studied. The duplexes were formed by a stem-loop and its complementary target. The thermal denaturations of the structures were monitored via circular dichroism (CD). Measurements of the kinetics of stem-loop and target hybridization were monitored by fluorescence resonance energy transfer (FRET, chapter 2.2). The selected donor and acceptor fluorophores were 5’6-carboxyfluorescein (FAM) and hexachlorofluorescein (HEX) respectively. FAM has an excitation maximum at 494 nm and an emission maximum at 520 nm. The excitation and emission maximum for HEX are at 538 nm and 555 nm respectively.
Scheme 4.1 – The cyclin D1 and survivin sequences. (A) Cyclin D1 stem-loop and target strands and (B) the survivin stem-loop and target strands. The underlined bases in (A) refer to the bases which are exclusive to the 36 base pair target. In both panels, bold denotes bases which are involved in duplex base pairing.
4.2 Materials and Methods

4.2.1 Materials

Chemicals

The salts, NaCl, KCl, Na₂HPO₄, NaH₂PO₄ and tris base were purchased from Sigma-Aldrich. The [ethylenedinitrilo]-tetraacetic acid disodium salt (EDTA) was purchased from Mallinckrodt AR. The boric acid and glycerol were purchased from EMD Biosciences. The Nonidet-P40 (NP40) was purchased from US Biological. Ethidium bromide (EtBr) was obtained from MP Biomedicals. The polymers, poly(ethylene glycol) (PEG) of 3400 (3.4K) and 8000 (8K) average molecular weight and glycerol were purchased from Sigma-Aldrich. The polymers were dissolved in phosphate buffered saline (PBS). The macrosolute concentrations are reported as weight per volume (w/v).

Oligonucleotides

The labeled and unlabeled oligonucleotides were obtained from Integrated DNA Technologies Inc (Coralville, IA). The oligonucleotides were dissolved in PBS (pH 7.4, 137 mM NaCl, 2.6 mM KCl, 10 mM Na₂HPO₄ and 1.8 mM NaH₂PO₄). The concentrations of the oligonucleotides were determined by ultra-violet (UV) absorbance on a Beckman spectrophotometer (model DU-640), using the extinction coefficients at 260 nm (Table 4.1). Complementary oligonucleotides were mixed in a 1:1 concentration ratio for the formation of duplex oligonucleotides. The oligonucleotides
were annealed by heating at 95 °C for four minutes after which they were cooled to room temperature over a period of two hours.

<table>
<thead>
<tr>
<th></th>
<th>$\varepsilon_{260\text{ nm}}$ Labeled ($M^{-1} cm^{-1}$)</th>
<th>$\varepsilon_{260\text{ nm}}$ Unlabeled ($M^{-1} cm^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cyclin D1 stem-loop</td>
<td>255,000</td>
<td>223,400</td>
</tr>
<tr>
<td>Cyclin D1 3’ Ext</td>
<td>293,800</td>
<td></td>
</tr>
<tr>
<td>Cyclin D1 5’ Ext</td>
<td>289,600</td>
<td></td>
</tr>
<tr>
<td>Cyclin D1 target-29</td>
<td>267,100</td>
<td></td>
</tr>
<tr>
<td>Cyclin D1 target-36</td>
<td>340,100</td>
<td></td>
</tr>
<tr>
<td>Survivin Stem-loop</td>
<td>274,900</td>
<td>243,300</td>
</tr>
<tr>
<td>Survivin target</td>
<td>257,000</td>
<td></td>
</tr>
</tbody>
</table>

Table 4.1 – The extinction coefficients (260 nm) for the cyclin D1 and survivin sequences. The extinction coefficients were obtained from IDT Inc. and were based on the salt concentrations of PBS buffer.

4.2.2 Measurements

Thermal Denaturation of the Stem-loops and Duplexes

Melt profiles were obtained in dilute and crowded environments. The dilute environment was created with a simple buffer (PBS). The crowded environment was created with the use of a macrosolute (20% PEG 8K). The thermal denaturation of the annealed duplexes was monitored via circular dichroism (CD) as the change in ellipticity at 280 nm on a Jasco J-710 spectropolarimeter. An external computer-controlled bath was used to control the temperature of the cell holder. All melts were
conducted in a 1.0 cm path length quartz cuvette. The samples were overlaid with mineral oil to prevent evaporation during the progress of the experiment. Prior to starting the melt, the samples were equilibrated for 15 minutes at the start temperature. Wavelength scans were collected from 200-300 nm at the beginning and end of each melt. Thermal melt profiles were obtained by collecting the ellipticity at 280 nm as the temperature increased at a rate of 20 °C per hour. All scans were background corrected against buffer (PBS) or crowded solution. A concentration of 10 µM was used for the stem-loops and each of the duplex strands. The values reported are based on the average of three or more individually fit trials. The thermal profiles were created by converting the change in ellipticity at 280 nm (Δε_{280}) into the fraction of secondary structure (α) and graphing α versus temperature.

**Kinetics of Hybridization of the Duplexes**

Electrophoretic mobility shift assays (EMSAs) were used to study the kinetics of hybridization of the stem-loops to their respective targets. The concentration was 5 nM for the HEX labeled stem-loops and 50 nM for the unlabeled target strands. The kinetics of hybridization were studied in dilute conditions (PBS), in the presence of macrosolutes (20% PEG 3.4K and 8K) and in the presence of an osmolyte (20% glycerol). The PBS samples were prepared in 1X reaction buffer (1mM EDTA, 0.1% NP40, 5% glycerol and 1X PBS). The crowded samples were prepared in reaction
buffer without glycerol. For the kinetics of the duplex formation in PBS and glycerol, samples were incubated prior to loading with incubation times ranging from 0-7 hours. For the kinetics of the duplex formation in 20% PEG 3.4K and 8K, samples were loaded in five minute increments onto a running gel. All samples were loaded onto 8% polyacrylamide gels (80:1 crosslink) and run for one hour at 300 volts in running buffer (22 mM Tris, 22.2 mM boric acid, 0.5 mM EDTA). The gels were visualized by utilizing HEX-labeled stem-loops. The gels were scanned and quantified on a Bio-Rad Molecular Imager (model FX). The intensity of all the bands were background corrected against the intrinsic intensity of the gel. Kinetics of hybridization profiles were created by graphing the fraction of duplexes versus time. The values reported are based on the fit of an average of three or more trials with a pseudo first order kinetic model.

4.2.3 Data Analysis

Analysis of the Stem-loop Melt Profiles

The melt profiles of the stem-loops were obtained by plotting the fraction of secondary structure versus the temperature in Kelvin (K). The melt profiles were fit with an expression for two-state melting of a self-complementary sequence with Mathematica 6.0. For a unimolecular system, where the molecularity (n) is equal to one, the $K_{eq}(T)$ expression is
\[
K_{eq}(T) = \frac{\alpha}{(C_T)^n(1-\alpha)^n} = \frac{\alpha}{(1-\alpha)} \quad \text{(Equation 4.1)}
\]

where \(C_T\) represents the total strand concentration and \(\alpha\) is the fraction of secondary structure of the stem-loop (23,24). The van’t Hoff expression for \(K_{eq}(T)\) is:

\[
K_{eq}(T) = \exp\left(-\frac{\Delta G}{RT}\right) = \exp\left(-\frac{\Delta H}{RT} + \frac{\Delta S}{R}\right) \quad \text{(Equation 4.2)}
\]

where \(\Delta G\) is the change in free energy, \(\Delta H\) is the change in enthalpy, \(\Delta S\) refers to the change in entropy and \(R\) is the gas constant, 1.987 cal mol\(^{-1}\) K\(^{-1}\) (23). Equation 4.1 is solved in terms of \(\alpha\), and the van’t Hoff expression for \(K_{eq}\) is substituted into this expression.

\[
\alpha(T) = \frac{\exp\left(-\frac{\Delta H}{RT} + \frac{\Delta S}{R}\right)}{1 + \exp\left(-\frac{\Delta H}{RT} + \frac{\Delta S}{R}\right)} \quad \text{(Equation 4.3)}
\]

The \(\alpha\) term defined in equation 4.3 was used to analyze the melting profiles for the cyclin D1 and survivin stem-loops. The \(\Delta H\) and \(\Delta S\) values were obtained from the curve fit analysis and were used to calculate the melting temperature (\(T_m\)). The following general equation for \(T_m\) can be applied to a self-complementary system of any molecularity (23).

\[
\frac{1}{T_m} = \frac{(n-1)R}{\Delta H} \ln C_T + \frac{[\Delta S-(n-1)R\ln(2)+R\ln(n)]}{\Delta H} \quad \text{(Equation 4.4)}
\]

The \(T_m\) expression for the stem-loops (\(n=1\)) can be defined as

\[
T_m = \frac{\Delta H}{\Delta S}. \quad \text{(Equation 4.5)}
\]
The Gibbs free energy ($\Delta G$), was calculated at 37 °C because that temperature is physiologically relevant (24).

$$\Delta G = \Delta H - T\Delta S = -RT\ln K_{eq} \quad \text{(Equation 4.6)}$$

The melt profiles were individually fit to obtain $\Delta H$, $\Delta S$ and calculate $\Delta G_{37^\circ C}$ and $T_m$ values. The reported $\Delta H$, $\Delta S$, $\Delta G_{37^\circ C}$ and $T_m$ are average values based on a minimum of 3 trials. The error for parameters was calculated as standard error of the mean.

Prior to curve-fitting, the change in ellipticity data for the stem-loops was converted into $\alpha$ using equation 4.7 (24).

$$\alpha = \frac{[\epsilon_{280}(T_{\max})+m_d(T-T_{\max})]-[\epsilon_{280}(T)]}{[\epsilon_{280}(T_{\max})+m_d(T-T_{\max})]-[\epsilon_{280}(T_{\min})-m_{nd}(T-T_{\min})]} \quad \text{(Equation 4.7)}$$

where $\epsilon_{280}(T_{\min})$ is the average measured ellipticity at the temperature minimum ($T_{\min}$) and $\epsilon_{280}(T_{\max})$ is the average measured ellipticity at the temperature maximum ($T_{\max}$) (24). The $m_d$ and $m_s$ are the average slopes for the first and last 10-15 °C temperature points respectively. The $m_{nd}$ represents the temperature dependence of the ellipticity of the non-denatured (nd) DNA and $m_d$ represents the temperature dependence of ellipticity the denatured (d) DNA (24). This procedure corrected for the temperature dependence of the upper and lower baselines.
Analysis of the Duplex Melt Profiles

The survivin and cyclin stem-loops were annealed with their respective target strands in order to study the effects of crowding on the duplex structures. The formation of the duplexes can be expressed as:

\[ A + B \rightleftharpoons AB \]  \hspace{1cm} (Equation 4.8)

For the duplex (n=2) the \( K_{eq}(T) \) for the association of the structure can be defined as:

\[
K_{eq}(T) = \frac{[AB]}{[A][B]} = \frac{\alpha}{(c_T)^n(1-\alpha)^n} = \frac{2\alpha}{(c_T)(1-\alpha)^2}. \hspace{1cm} (Equation 4.9)
\]

Mathematica 6.0 was used to solve equation 4.9 in terms of \( \alpha \). The negative form of the root was real in the range of the experimental data. The van’t Hoff expression for \( K_{eq} \) was then substituted into the expression for \( \alpha \).

\[
\alpha(T) = \frac{1+C_T \exp\left(-\frac{\Delta H}{RT} + \frac{\Delta S}{R}\right) - \sqrt{1+2C_T \exp\left(-\frac{\Delta H}{RT} + \frac{\Delta S}{R}\right)}}{C_T \exp\left(-\frac{\Delta H}{RT} + \frac{\Delta S}{R}\right)} \hspace{1cm} (Equation 4.10)
\]

This expression was then used to analyze the melt data for the duplexes. The following general equation for \( T_m \) can be applied to a non-self complementary system of any molecularity (23).

\[
\frac{1}{T_m} = \frac{(n-1)R}{\Delta H} \ln C_T + \frac{[\Delta S-(n-1)R\ln(2n)]}{\Delta H} \hspace{1cm} (Equation 4.11)
\]

Hence, for the duplex the \( T_m \) expression was defined as:
\[ T_m = \frac{\Delta H}{\Delta S + R \ln \left( \frac{C_T}{4} \right)} \quad \text{(Equation 4.12)} \]

The \( \Delta G_{37c} \) and its error were calculated by the same method used for the stem-loop. The conversion of the data prior to curve fitting was also conducted by the same method used for the stem-loop data.

**Analysis of the Hybridization Profiles for CyclinD1 and Survivin**

The hybridization profiles of cyclin D1 and survivin were fit using Kaleidagraph 3.6. The bands on the gels were quantified using Quantity One software. The band intensities were converted into fraction of duplex (\( F_D \)), and were fit with a pseudo first order kinetic model.

\[ F_D = 1 - (f_1 \exp^{-kt} + f_2) \quad \text{(Equation 4.13)} \]

In equation 4.13, \( f_1 \) and \( f_2 \) correct for the range and minimum of the data, \( t \) is time in seconds, and \( k \) is the pseudo first rate (s\(^{-1}\)) which is composed of [target]\(*k\) where \( k \) is the second order rate constant.
4.3 Results

4.3.1 Macrosolute Effects on the Stem-loops and Duplexes

Melt profiles of the cyclin D1 and survivin stem-loops were conducted in a dilute environment created by a simple buffer (PBS), while crowded conditions were created by the presence of a macrosolute, 20% PEG 8K. Macrosolute selection was based on previous effects noted with other nucleic acid systems (chapter 2) and published reports of duplex stabilization in the presence of 20% PEG 8K (12,15,18). The thermodynamic stabilities of the stem-loops were studied by heating the structures and monitoring the loss of secondary structure via CD at 280 nm. Figures 4.1 and 4.2 contain representative wavelength scans for cyclin D1 and survivin stem-loops respectively. Wavelength scans were collected at the start and end of a melt. Based on the wavelength scans, it is evident that the unimolecular structures are B-form as indicated by the characteristic peak minimum (245 nm) and maximum (280 nm) (25). The monitoring wavelength (280 nm) was selected because there is a large change in signal at that wavelength when the structures are completely denatured.

Figure 4.1 contains the individual melt profiles for the cyclin D1 stem-loop in PBS and in the presence of 20% PEG 8K. The melting point of a structure (T_m) is the temperature at which half of the secondary structure is lost. The T_m of the cyclin D1 stem-loop in dilute environment was 56.5 °C and increased to 57.9 °C in the presence of macrosolute. Hence, the difference in T_m between the crowded as compared to the dilute environment (ΔT_m) was 1.3 °C (Table 4.2). An additional
comparison that can be made is between the $\Delta G_{37}^\circ$ in crowded and dilute environments, $\Delta\Delta G_{37}^\circ$. The $\Delta\Delta G_{37}^\circ$ is a thermodynamic parameter that is commonly used as a standard to compare the differences in stability between different structures. The $\Delta\Delta G_{37}^\circ$ for the cyclin D1 stem-loop is -0.1 kcal mol$^{-1}$. This $\Delta\Delta G_{37}^\circ$ value indicates that although the $\Delta T_m$ was 1.3 °C for the cyclin D1 stem-loop, no significant stabilization was noted in the presence of the macrosolute.

Representative wavelength scans of survivin depict the B-form nature of the unimolecular structure and the significant decrease in signal at 280 nm upon total denaturation (Figure 4.2). The $T_m$ of a stem-loop is determined by the number of base pairs that comprise the stem. The cyclin D1 stem-loop has a six base pair stem and a $T_m$ of 56.6 °C in PBS. The $T_m$ of the five base pair survivin stem-loop is 46.9 °C in PBS (Table 4.2). The $G_{37}^\circ$ of survivin stem-loop in PBS was -0.9 kcal mol$^{-1}$ while the $G_{37}^\circ$ of cyclin D1 in PBS was -2.7 kcal mol$^{-1}$. Hence, the cyclin D1 stem-loop was 1.8 kcal mol$^{-1}$ more stable than survivin in PBS.

In the presence of 20% PEG 8K, the $T_m$ of survivin increased to 49.2 °C, a $\Delta T_m$ of 3.2 °C. While the stabilization may appear significant, the comparison of the $\Delta\Delta G_{37}^\circ$ demonstrates the stabilization is marginal with a $\Delta\Delta G_{37}^\circ$ of -0.2 kcal mol$^{-1}$. Hence, although the stabilities of the two stem-loops are quite different, the crowding mediated stabilization of the two structures is similar.
Figure 4.1 – Representative wavelength scans and the melt profiles of the cyclin D1 stem-loop in PBS and 20% PEG 8K. Scans were conducted at the start (blue) and end (red) of melts. The melt profiles in PBS and PEG were curve fit as depicted by the solid black and blue lines respectively. (A) Cyclin D1 stem-loop scans at 20 and 90 °C in PBS (B) cyclin D1 stem-loop melts in PBS (C) cyclin D1 stem-loop scans at 15 and 90 °C in 20% PEG 8K (D) cyclin D1 stem-loop melts in 20% PEG 8K.
Figure 4.2 – Representative wavelength scans and the melt profiles of the survivin stem-loop in PBS and 20% PEG 8K. Scans were conducted at the start (blue) and end (red) of melts. The melt profiles in PBS and PEG were curve fit as depicted by the solid black and blue lines respectively. (A) Survivin stem-loop scans at 15 and 95 °C in PBS (B) survivin stem-loop melts in PBS (C) survivin stem-loop scans at 15 and 95 °C in 20% PEG 8K (D) survivin stem-loop melts in 20% PEG 8K.
<table>
<thead>
<tr>
<th></th>
<th>Cyclin D1 stem-loop</th>
<th>Cyclin D1 Duplex</th>
<th>Survivin stem-loop</th>
<th>Survivin Duplex</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>ΔH (kcal mol⁻¹)</strong></td>
<td>-45.0 ± 0.9</td>
<td>43.8 ± 1.0</td>
<td>-74.6 ± 0.1</td>
<td>-80.8 ± 1.3</td>
</tr>
<tr>
<td><strong>ΔS (cal mol⁻¹ K⁻¹)</strong></td>
<td>-136.5 ± 2.9</td>
<td>-132.3 ± 3.4</td>
<td>-194.4 ± 0.1</td>
<td>-210.5 ± 3.4</td>
</tr>
<tr>
<td><strong>T_m (°C)</strong></td>
<td>56.6 ± 0.5</td>
<td>58.1 ± 0.8</td>
<td>65.8 ± 0.4</td>
<td>69.0 ± 0.6</td>
</tr>
<tr>
<td><strong>ΔG₃⁷°C (kcal mol⁻¹)</strong></td>
<td>-2.7 ± 0.1</td>
<td>-2.8 ± 0.4</td>
<td>-14.3 ± 0.1</td>
<td>-15.5 ± 0.2</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>PBS</th>
<th>20% PEG 8K</th>
<th>Δ PEG 8K-PBS</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>ΔH (kcal mol⁻¹)</strong></td>
<td>-45.0 ± 0.9</td>
<td>43.8 ± 1.0</td>
<td>-74.6 ± 0.1</td>
</tr>
<tr>
<td><strong>ΔS (cal mol⁻¹ K⁻¹)</strong></td>
<td>-136.5 ± 2.9</td>
<td>-132.3 ± 3.4</td>
<td>-194.4 ± 0.1</td>
</tr>
<tr>
<td><strong>T_m (°C)</strong></td>
<td>56.6 ± 0.5</td>
<td>58.1 ± 0.8</td>
<td>65.8 ± 0.4</td>
</tr>
<tr>
<td><strong>ΔG₃⁷°C (kcal mol⁻¹)</strong></td>
<td>-2.7 ± 0.1</td>
<td>-2.8 ± 0.4</td>
<td>-14.3 ± 0.1</td>
</tr>
</tbody>
</table>

Table 4.2 – The thermodynamic parameters of hybridization for the cyclin D1 and survivin stem-loops and duplexes. All values were determined from the thermal denaturation studies of the structures in PBS, and in the presence of 20% PEG 8K. Reported values are the average of three or more individually fit trials. Reported errors were calculated as the standard error of the mean. The ΔT_m and ΔΔG₃⁷°C errors were tabulated using propagation of error.

a denotes that the values obtained are not reliable
The Crowding Effects on the Cyclin D1 and Survivin Duplexes

The cyclin D1 stem-loop and the target-36 strands comprise the cyclin D1 duplex (Scheme 4.1). Thermal denaturation of the duplex was monitored by the change in ellipticity at 280 nm in the presence of 20% PEG 8K and in dilute buffer (PBS). Figure 4.3 shows representative wavelength scans and individual melt profiles for the B-form cyclin D1 duplex. The T_m of the 17 base pair cyclin D1 duplex was 65.8 °C in PBS. In comparison to the cyclin D1 stem-loop, the duplex structure is more stable, as shown by the duplex ΔG_{37C} of -16.6 kcal mol^{-1} in PBS. In the presence of PEG 8K, the T_m of the duplex increased by 3.2 °C. The increase in T_m is significant as confirmed by the ΔΔG_{37C} of -1.4 kcal mol^{-1}.

The survivin duplex was composed of the survivin stem-loop and target strands (Scheme 4.1). Wavelength scans for the survivin duplex, and a large change in signal at 280 nm upon thermal denaturation, indicate B-form structure (Figure 4.4). The survivin duplex was thermally denatured in PBS and in the presence of 20% PEG 8K. The survivin duplex appears to be stabilized in the presence of 20% PEG 8K. However, although the melt profiles of the survivin data were reproducible and consistent with one another, a reliable fit through the data was not possible. The upper baseline contained a temperature dependence which could not be corrected for as shown in Figure 4.4.
Figure 4.3 – Representative wavelength scans and the melt profiles for the cyclin D1 duplex in PBS and 20% PEG 8K. Scans were conducted at the start (blue) and end (red) of melts. The melt profiles in PBS and PEG were curve fit as depicted by the solid black and blue lines respectively. (A) Cyclin D1 duplex scans at 20 and 90 °C in PBS (B) cyclin D1 duplex melts in PBS (C) cyclin D1 duplex scans at 20 and 90 °C in 20% PEG 8K (D) cyclin D1 duplex melts in 20% PEG 8K.
Figure 4.4 – Representative wavelength scans and the melt profiles for the survivin duplex in PBS and 20% PEG 8K. Scans were conducted at the start (blue) and end (red) of melts. The melt profiles in PBS and PEG were curve fit as depicted by the solid black and blue lines respectively. (A) Survivin duplex scans at 20 and 90 °C in PBS (B) survivin duplex melts in PBS (C) survivin duplex scans at 20 and 90 °C in 20% PEG 8K (B) survivin duplex melts in 20% PEG 8K.
4.3.2 Macrosolute Effects on Kinetics of Hybridization of the Duplexes

In addition to stabilizing structures, macrosolutes can enhance the kinetics of hybridization (26). Although the introduction of a macrosolute increases the viscosity of a system, the increase in chemical activity opposes viscosity effects (26,27). Specifically, the presence of macrosolutes can result in an increase in the forward rate of a reaction and subsequent decrease in the reverse rate (26). The kinetics of duplex formation between the cyclin D1 and survivin stem-loop and their respective target strands was monitored by EMSA. The dilute environment was created with PBS, and the crowded environments were mimicked using the macrosolutes (20% PEG 3.4K or 8K), and an osmolyte (20% glycerol). The stem-loops were labeled with HEX which allowed for visualization of the gel. For both systems, the bound stem-loop migrated less than the free stem-loop (figures 4.5 and 4.7). As time progressed, the bound stem-loop band increased in intensity while the intensity of the free stem-loop band decreased (figures 4.5 and 4.6). The intensities of the bands were quantified and converted to fraction of duplex and fit with a pseudo first order kinetic model.
Figure 4.5 – Representative EMSA for the kinetics of hybridization of the cyclin D1 stem-loop to the target-36 strand. (A) PBS (B) 20% glycerol (C) 20% PEG 3.4K and (D) 20% PEG 8K. The 5 nM HEX labeled cyclin stem-loop and 50 nM target-36 strands were mixed and then loaded onto 8% polyacrylamide gels after the indicated incubation times. For the 20% PEG 3.4K and 8K experiments, the samples were loaded on a running gel. For all cyclin D1 gels, the identity of the middle band was concluded to be a stem-loop homoduplex (Figure 4.6).

The cyclin D1 kinetics of duplex formation gels contained a total of three bands per time point which in order of decreasing migration were; free stem-loop, an unknown complex band and the stem-loop:target duplex (Figure 4.5). The identity of the unknown complex band was probed by changing the length of the target strand and the length of the stem-loop. The modified target strand, target-29, was identical to target-
36 but with the first seven 5’ end bases removed. If the unknown complex band involved the target strand, the mobility of the band would change with the truncation of the target strand. Additionally, two extended, unlabeled, cyclin D1 stem-loops cyclin D1 5’ ext and cyclin D1 3’ ext (Scheme 4.2) were created. The extended stem-loops contained an additional 7 bases on the 5’ or 3’ end. The additional bases did not alter the number of base pairs in the stem, but rather elongated the strand and thus would change the mobility of the unknown complex band if it was a structure involving only stem-loop.

**Scheme 4.2 - The cyclin D1 stem-loop 5’ and 3’ extended sequences.** The bold bases are complementary to the target strands. The underlined bases denote the seven bases added to extend the length of the strands.

EMSA analysis of the extended cyclin D1 stem-loops and truncated target strands is shown in Figure 4.6. The unlabeled samples were prepared in 5% PEG 8K and were incubated for 1 hour prior to loading on the gel. Ethidium bromide (EtBr) staining was used for visualization of the gel since all of the samples were unlabeled. As expected, the migration of the original cyclin stem-loop was faster than that of the extended stem-loops (lanes 1-3). Similarly the migration of the target-29 strand was
faster than target-36 (lanes 11 and 12). The unknown complex band did not demonstrate a change in migration when the cyclin D1 stem-loop was incubated with target-29 versus target-36. However, the duplex band exhibited a faster migration pattern with the shorter target (lane 4 versus lane 7). The migration of the unknown complex changed with the extended stem-loops. When target-29 was incubated with the cyclin D1 stem-loop versus the extended stem-loops, the migration of the unknown complex band decreased (lanes 4-6). Similarly, when target-36 was incubated with the different stem-loops the migration of the complex band decreased for the extended versus cyclin D1 stem-loop (lanes 7-9). Therefore, the identity of the additional band on the cyclin D1 kinetic gels is most likely cyclin D1 stem-loop homoduplex.

The homoduplex is a 12 base pair structure in which the bases in the stem of one cyclin D1 strand form base pairs with the stem bases on another strand. The two 6 base pair regions are separated by the bases of the loop. The stem-loop homoduplex is present only in kinetic reactions and is absent in the annealed product. Thus, the stem-loop homoduplex is a kinetic intermediate. It is possible that the overhang regions which result upon stem-loop:target hybridization, make the system more susceptible to opening of stem-loops and formation of stem-loop dimmers.
Figure 4.6 – Identification of the unknown complex band in the cyclin D1 kinetic studies. All samples were prepared in 5% PEG 8K and incubated for 1 hour prior to loading on the 8% polyacrylamide (80:1 crosslink) gels. Visualization of the gel was done with EtBr staining. In lanes 1-3 and 11-12, the strand concentration was 400 nM per strand. In lanes 4-9, stem-loop and target strand concentration was 150 nM per strand. In lane 10 stem-loop strand concentration was 150 nM and target concentration was 75 nM. (1) cyclin D1 stem-loop (2) cyclin D1 3’ ext (3) cyclin D1 5’ ext (4) cyclin D1 stem-loop and target-29 (5) cyclin D1 3’ ext and target-29 (6) cyclin D1 5’ ext and target-29 (7) cyclin D1 stem-loop and target-36 (8) cyclin D1 3’ ext and target-36 (9) cyclin D1 5’ ext and 150 nM target-36 (10) cyclin D1 stem-loop, target-29 and target 36 (11) target-29 (12) target-36.

As a result of the stem-loop homoduplex formation in the cyclin D1 kinetic gels, the kinetics of duplex formation (stem-loop:target) were analyzed in terms of the duplex band for both the cyclin D1 and survivin studies. In the dilute environment, the cyclin D1 duplex had a rate constant of $2.3 \times 10^4 \text{ M}^{-1}\text{s}^{-1}$ in PBS (Figure 4.8 and Table 4.3). In the presence of the macrosolute 20% PEG 3.4K the rate constant increased by 9.5-fold. A larger enhancement in rate constant was exhibited in the presence of the macrosolute 20% PEG 8K with over a 12-fold increase. The osmolyte used in this study, 20% glycerol, resulted in slight increase in the rate constant (1.0-fold).
When the cyclin D1 kinetics were analyzed in terms of the intensity of the stem-loop homoduplex band, the rate constants were smaller than the rate constants obtained in terms of the stem-loop:target duplex band. In PBS and glycerol the rate constant of stem-loop homoduplex formation was 1.4 and 1.0 M$^{-1}$s$^{-1}$ respectively. The presence of glycerol resulted in a slight decrease of the rate constant by 0.74-fold. In the presence of 20% PEG 3.4K and 8K the rate constant increased by 9.5 and 11.9-fold respectively. Hence, the kinetics of stem-loop:target duplex formation and the kinetics of stem-loop homoduplex were enhanced to similar extents in the presence of the macrosolutes.

EMSA analysis of the kinetics of survivin duplex formation was simpler than that of cyclin D1 because only two bands (free and bound (duplex) stem-loop) were present on the gel (Figure 4.7). In the dilute environment formation of the survivin duplex had a rate constant of $1.7 \times 10^4$ M$^{-1}$s$^{-1}$ (Figure 4.8 and Table 4.3). In the presence of the macrosolutes 20% PEG 3.4K and PEG 8K, the rate constant of duplex formation increased by 5.5 and 5.6-fold respectively. In the presence of the osmolyte, 20% glycerol, the rate constant of duplex formation for survivin increased slightly (1.1-fold).
Figure 4.7 – Representative EMSA for the kinetics of hybridization of the survivin stem-loop to the target strand. (A) PBS (B) 20% glycerol (C) 20% PEG 3.4K and (D) 20% PEG 8K. The 5 nM HEX labeled survivin stem-loop and 50 nM target strands were mixed and then loaded onto 8% polyacrylamide gels after the indicated incubation times. For the 20% PEG 3.4K and 8K experiments, the samples were loaded on a running gel. The * denotes a survivin stem-loop and target which were incubated for two weeks. The D denotes a previously annealed duplex and SL denotes stem-loop in the absence of target.
Figure 4.8 – Kinetics of the cyclin D1 and survivin hybridization reactions in the presence of various cosolutes. The kinetics of duplex formation were conducted in PBS (black), in the presence of 20% PEG 3.4K (blue), PEG 8K (red) and glycerol (gray). (A) The kinetics of duplex formation for the cyclin D1 stem-loop and target-36 strands and (B) the kinetics of duplex formation for the survivin stem-loop and target strands. Errors bars depict standard error obtained from the average of at least three trials.

<table>
<thead>
<tr>
<th></th>
<th>Cyclin K (10^4 M⁻¹ s⁺¹)</th>
<th>Survivin K (10^4 M⁻¹ s⁺¹)</th>
</tr>
</thead>
<tbody>
<tr>
<td>PBS</td>
<td>2.3 ± 0.3</td>
<td>1.7 ± 0.3</td>
</tr>
<tr>
<td>20% PEG 3.4K</td>
<td>21.9 ± 1.5</td>
<td>9.4 ± 0.9</td>
</tr>
<tr>
<td>20% PEG 8K</td>
<td>28.7 ± 2.6</td>
<td>9.5 ± 1.5</td>
</tr>
<tr>
<td>20% Glycerol</td>
<td>2.4 ± 0.3</td>
<td>1.9 ± 0.4</td>
</tr>
</tbody>
</table>

Table 4.3 – The rate constants of formation for cyclin D1 and survivin duplexes in the presence of various cosolutes. Values are based on the average of three or more trials. Reported errors are the standard error of predication from the fit.
4.4 Discussion

4.4.1 Macrosolutes Marginally Enhance the Stability of Stem-loops

The cyclin D1 and survivin stem-loops are both unimolecular structures which were marginally stabilized in the presence of the macrosolute. In the presence of 20% PEG 8K the $\Delta \Delta G_{37^\circ C}$ of cyclin D1 and survivin stem-loops were -0.1 and 0.2 kcal mol$^{-1}$ respectively. However, the stability of the cyclin D1 duplex increased under the crowded conditions as indicated by the $\Delta \Delta G_{37^\circ C}$ of -1.4 kcal mol$^{-1}$. Hence, the cyclin D1 duplex was stabilized over the stem-loop by -1.3 kcal mol$^{-1}$. The average reported value for duplex $T_m$ increase is 3-5 °C in the presence of 20% PEG 8K (15,18,28). Hence, the macrosolute induced increase in $T_m$ for the cyclin D1 duplex of 3.2 °C correlates well with previous reports. Therefore, as the molecularity decreased we noted a decrease in crowding effects.

Macromolecular crowding will favor the more compact form of a macromolecule, since the more compact form will increase the available volume available to cosolutes (29). Hence, with the appropriate cosolute, a duplex will be favored over single strands in a crowded environment since its co-volume is smaller than that of the single strands (28). The duplexes studied here differed from other duplexes previously studied (chapters 2 and 3). The cyclin D1 and survivin duplexes both contained overhang sequence(s) (Scheme 4.1) which may have resulted in a lower reduction in co-volume as compared to the single strands. Thus, the crowding mediated stabilization may have been reduced as a result of the overhangs.
4.4.2 Macrosolute Effects Correlate With Increased Activation Energy

The kinetics of the stem-loop and target hybridization reactions were studied in crowded and dilute environments via EMSA. The hybridization reactions presented here are more complex than those previously studied (chapters 2 and 3) where the duplex formation occurs when two non-self complementary strands hybridize. In these studies, one of the strands which comprised the duplex has secondary structure because it is a stem-loop. During or prior to hybridization to the target, the stem-loop must lose its structure and open. The presence of the macrosolutes enhanced the rate constant of formation of survivin and cyclin D1 duplexes. The typical rate constant for duplex formation is \(10^5\) to \(10^7\) M\(^{-1}\)s\(^{-1}\) (30). The rate constants of formation in these studies are slower which suggest an increased activation energy barrier for the hybridization reactions.

In previous studies, we obtained a 3-fold increase in the rate constant of hybridization for the BRCA1-185 16 base pair duplexes in the presence of 20% PEG 8K (chapter 2). The survivin duplex was enhanced by 5.6-fold in the presence of 20% PEG 8K and cyclin D1 duplex formation was enhanced by over an order of magnitude (12.4-fold) in the presence of 20% PEG 8K. We conclude that the macrosolute effects on the kinetics of formation, may increase with increasing activation energy of a system.

In the presence of 20% PEG 3.4K and 8K a significant increase in the rate constant of formation versus the dilute environment was observed for both duplexes.
However, the crowding-mediated enhancement of the rate constant was greater for cyclin D1. As the size of the macrosolute increased from PEG 3.4K to P8K, the effects on the cyclin D1 rate constant also increased. However, a similar trend was not observed for survivin whose rate constant of duplex formation was enhanced to similar extents in both PEG 3.4K and 8K.

Previous work on the kinetics of duplex formation confirmed that the macrosolutes led to enhancement of the rate constant while the presence of osmolytes, such as glycerol, led to a reduction in the rate constant (chapter 2). However, cyclin D1 and survivin duplexes exhibited a slight increase in the rate constant in the presence of glycerol. The marginal enhancement effect in the presence of glycerol is likely due to the self-complementarity of a stem-loop, which requires an open conformation prior to hybridization. We propose that the stem-loop structure was destabilized by the presence of glycerol, which resulted in a slight increase of the rate constant of duplex formation. The unexpected behavior of the hybridization reaction in glycerol, and the formation of the stem-loop homoduplex with cyclin D1, highlights the complexity of stem-loop:target reactions. Although a stem-loop:target duplex is the expected kinetic product in these hybridization reactions, it is possible to form alternate structures as demonstrated by the cyclin D1 homoduplex.
4.5 Conclusions

The work on the cyclin D1 and survivin structures demonstrated that macrosolutes can result in thermal stability enhancement of duplex structures but only marginal stabilized the stem-loop structures. The crowding mediated stabilization decreased as the molecularity of the structures decreased. A substantial crowding enhancement of the rates of duplex formation between the stem-loops and their respective target strands was observed in the presence of macrosolutes. In the presence of glycerol, a known reducer of duplex rate of formation, a slight increase in the rate constant was noted. These results can be explained by the glycerol destabilizing the stem-loop structure, and facilitating the formation of a duplex. In the presence of PEG 8K, the cyclin D1 duplex exhibited over an order of magnitude increase in the rate constant of formation. Comparison of the observed cyclin D1 crowding-induced kinetic effects with previously studied duplexes leads to the conclusion that as the activation energy of the systems increases, crowding effects also increase. These results are of importance to molecular beacon assays or other techniques which rely on nucleic acid hybridization. The crowding effects on the cyclin D1 and survivin systems demonstrate that macromolecular crowding can result in consequences to thermodynamic and equilibrium values for stem-loop and duplex structures.
References


Chapter 5 - Macrosolute Effects on Multi-branch DNA Junctions

5.1 Introduction

Creating a simplified model of the intracellular environment with the use of cosolutes, inert, synthetic substitutes of cellular macromolecules, requires consideration of several factors (1). Some of the cosolute selection criteria include the following; first, the cosolute should not interact with the macromolecule being studied except for steric repulsion interactions (2). Second, the shape of the cosolute must be considered. One such example is the greater crowding effects can result when using globular versus rod-like shaped cosolutes (3). Third, the size of the cosolute is also an important factor. Nucleic acids systems in a crowded environment can exhibit enhanced stabilization or association rates when they are in the presence of macrosolutes (8-11). However, the use of smaller cosolutes, osmolytes, can result in the destabilization of nucleic acids and a decrease in association rates (4-6). In addition to cosolute size and shape, the macromolecule studied, also known as the test species can influence the crowding effects. For example, although osmolytes generally destabilize nucleic acids, it has been demonstrated that proteins can be stabilized against unfolding by the presence of osmolytes (7). The molecularity (n), the number of units that comprise the final structure of the test species, must also be considered as it can influence crowding mediated effects (8). One such example is that of a DNA duplex that exhibited a 5 °C increase in \( T_m \) in the presence of polyethylene glycol of average molecular weight 8000 (3,4). In the same study, a triplex structure had a 15 °C increase in \( T_m \).
In addition to linear duplex or triplex DNA, nucleic acids can exit in a variety of conformations such as stem-loops, quadruplex and multi-branch DNA (9,10). Unlike triplex DNA, multi-branch DNA does not have sequence or pH constraints of Hoogsteen base pairing to achieve stable structures (11,12). Multi-branch DNA structures can be formed from three or more complementary sequences and have been implicated in a variety of biological processes (11,13,14). These structures are present in Holliday junctions and have been proposed as intermediates in recombination and repair (15). Helical junction structures are also found in various RNA structures such as tRNA and ribozymes (16,17). Thermodynamic work on multi-branch junctions has resulted in valuable information on topics such as arm to arm distances, effects of mismatches and overall shapes of the structures (18-21). Studies on multi-branch DNA have also demonstrated the sensitivity of these structures to their environment. For example, it was demonstrated that the 4-way junction required high concentrations of Mg$^{2+}$ to form a Holliday junction (11). Multi-branch junctions have also been used as scaffolds for artificial DNA based nanostructures, such as DNA dendrimers and tetrahedrons (22,23). Despite their biological relevance and potential use in nanostructures, multi-branch structures have only been studied in dilute conditions (24-27). Here we present the effects of crowding on the thermodynamic and kinetic properties of sequence related structures whose molecularity increases from 2 to 4.
5.2 Materials and Methods

5.2.1 Materials

Chemicals

The salts, MgCl₂, NaCl, KCl, Na₂HPO₄, NaH₂PO₄ and tris base were purchased from Sigma-Aldrich. The [ethylenedinitrilo]-tetraacetic acid disodium salt (EDTA) was purchased from Mallinckrodt AR. The boric acid and glycerol were purchased from EMD Biosciences. The Nonidet-P40 (NP40) was purchased from US Biological. The ethidium bromide (EtBr) was obtained from MP Biomedicals. The macrosolute, poly(ethylene glycol) of average molecular weight 8000 (PEG 8K) was purchased from Sigma-Aldrich and dissolved in phosphate buffered saline (PBS). The PEG 8K concentration is reported as weight per volume (w/v).

Oligonucleotides

The unlabeled oligonucleotides were obtained from Integrated DNA Technologies Inc (Coralville, IA). The oligonucleotides were dissolved in PBS (pH 7.4, 137 mM NaCl, 2.6 mM KCl, 10 mM Na₂HPO₄ and 1.8 mM KH₂PO₄). The concentrations of the oligonucleotides were determined by ultra-violet (UV) absorbance on a Beckman spectrophotometer (model DU-640) using the extinction coefficients at 260 nm (Table 5.1). Complementary oligonucleotides were mixed in a 1:1 concentration ratio for the formation of duplex oligonucleotides. The oligonucleotides
were annealed by heating at 95 °C for four minutes after which they were cooled to room temperature over a period of two hours.

<table>
<thead>
<tr>
<th></th>
<th>$\varepsilon_{260,\text{nm}}$ (M$^{-1}$ cm$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>R$_c$ strand</td>
<td>197,800</td>
</tr>
<tr>
<td>R strand</td>
<td>174,600</td>
</tr>
<tr>
<td>H strand</td>
<td>190,500</td>
</tr>
<tr>
<td>X strand</td>
<td>190,800</td>
</tr>
<tr>
<td>R* strand</td>
<td>212,600</td>
</tr>
<tr>
<td>X$_4$ strand</td>
<td>192,100</td>
</tr>
<tr>
<td>B strand</td>
<td>184,600</td>
</tr>
</tbody>
</table>

Table 5.1– The extinction coefficients (260 nm) for the multi-branch sequences. The extinction coefficients were obtained from IDT Inc. and were based on the salt concentrations of PBS buffer.

5.2.2 Measurements

Thermal Denaturation of the Multi-branch Junctions

Melt profiles were obtained in PBS (dilute environment) and in the presence of 20% PEG 8K (crowded environment). The thermal denaturation of the annealed structures was monitored by circular dichroism (CD) on a Jasco J-710 spectropolarimeter. An external computer-controlled bath was used to control the temperature of the cell holder. All melts were conducted in a 1.0 cm path length quartz cuvette. The sample was overlaid with mineral oil to prevent evaporation during the
progress of the experiment. Prior to starting the melt, the samples were equilibrated for at least 15 minutes at the start temperature. Wavelength scans were collected from 200-300 nm before and after each melt. Thermal melt profiles were obtained by monitoring the change in ellipticity at 270 nm as the temperature increased at a rate of 20 °C per hour. All scans were background corrected against buffer (PBS) or crowded solution. The concentrations used were 2.5 µM per strand for the 3-arm structures and the RRc duplex, and 1.5 µM per strand for the 4-arm structure. The values reported are based on the average of three or more individually fit trials.

**Kinetics of Hybridization of RHX**

Hybridization profiles were monitored via CD at 270 nm at a constant temperature of 25 °C in PBS (dilute environment) and in the presence 20% PEG 8K (crowded environment). An external computer-controlled bath maintained the temperature of the cuvette holder. All scans were conducted in a 0.3 cm path length quartz cuvette. The samples were overlaid with mineral oil to prevent evaporation during the progress of the experiment. Prior to starting the experiment, the samples were individually equilibrated at 25 °C for 15 minutes. Upon completion of the equilibration period, the samples were mixed and data collection was initiated. A time-based CD spectrum at 270 nm was collected with a data density of 1 second for up to 1 hour. All scans were background corrected against buffer (PBS) or crowded solution. The concentration was 6.5 µM per strand for the kinetic studies of the formation of
RHX. All values reported are based on the average of three or more trials. The hybridization profiles were created by plotting the change in ellipticity at 270 nm versus time.

**Confirmation of the Multi-branch Structures**

Electrophoretic mobility shift assays (EMSAs) were performed in order to confirm the formation of the RHX, R*HX, RHX₄B and RRₑ structures. The concentration was 6.5 µM per strand for all structures. The PBS samples were prepared in 1X reaction buffer (1 mM EDTA, 0.1% NP40, 5% glycerol and 1X PBS). The crowded samples were prepared in reaction buffer without glycerol. The samples were incubated and then loaded onto 8% polyacrylamide gels (80:1 crosslink) and run for 75 minutes at 200 volts in running buffer (22 mM Tris, 22.2 mM boric acid, 0.5 mM EDTA and 133 mM NaCl). The gels were stained for one hour in 4 µg ml⁻¹ EtBr and imaged on a BioRad Molecular Imager (model FX).
5.2.3 Data Analysis

Analysis of the RHX and R*HX Melt Profiles

The melt profiles of RHX and R*HX were obtained by plotting the change in ellipticity at 270 nm, $\Delta \varepsilon_{270}$, versus temperature in Kelvin (K). The melt profiles were fit with an expression for two-state melting of a non-self-complementary sequence with Mathematica 6.0. For the RHX sequences, the reaction can be expressed as:

$$R + H + X \rightleftharpoons \text{RHX}.$$  \hspace{1cm} (Equation 5.1)

The molecularity ($n$) is the number of strands which comprise the final structure. Thus, for both RHX and R*HX, $n$ is three. The equilibrium constant ($K_{eq}(T)$), for the association of the 3-arm junction can be expressed as:

$$K_{eq}(T) = \frac{[\text{RHX}]}{[R][H][X]} = \frac{\alpha}{\left(\frac{C_T}{n}\right)^{n-1} (1-\alpha)^n} = \frac{9\alpha}{(C_T)^2 (1-\alpha)^3}$$ \hspace{1cm} (Equation 5.2)

where $C_T$ represents the total strand concentration and $\alpha$ is the fraction of 3-arm DNA (28,29). The van’t Hoff expression for $K_{eq}(T)$ is:

$$K_{eq}(T) = \exp\left(-\frac{\Delta G}{RT}\right) = \exp\left(-\frac{\Delta H}{RT} + \frac{\Delta S}{R}\right)$$ \hspace{1cm} (Equation 5.3)

where $\Delta G$ is the change in free energy, $\Delta H$ is the change in enthalpy, $\Delta S$ refers to the change in entropy and $R$ is the gas constant (28). Mathematica 6.0 was used to solve equation 5.2 in terms of $\alpha$, which yielded only one real root. Substitution of the van’t
Hoff expression for \( K_{eq} \) into the real root yields an expression for \( \alpha \) in terms of \( \Delta H \) and \( \Delta S \).

\[
\alpha = 1 + \frac{2^{1/3}}{3^{2/3}} \left( \frac{3C_T^4 \left( \exp \left( \frac{-\Delta H + \Delta S}{RT} \right) \right)^2 + \sqrt[3]{4C_T^6 \left( \exp \left( \frac{-\Delta H + \Delta S}{RT} \right) \right)^3 + 3C_T^8 \left( \exp \left( \frac{-\Delta H + \Delta S}{RT} \right) \right)^4}}{C_T^2 \left( \exp \left( \frac{-\Delta H + \Delta S}{RT} \right) \right)} \right)^{1/3}
\]

(Equation 5.4)

The change in ellipticity at 270 nm at a given temperature, \( \Delta \varepsilon_{270}(T) \) can be defined by the following expression

\[
\Delta \varepsilon_{270}(T) = \alpha [\varepsilon_{270}(T_{min}) + m_{nd}(T - T_{min})] + (1 - \alpha) [\varepsilon_{270}(T_{max}) - m_{d}(T_{max} - T)]
\]

(Equation 5.5)

where \( \varepsilon_{270}(T_{min}) \) is the average measured ellipticity at the temperature minimum (\( T_{min} \)) and \( \varepsilon_{270}(T_{max}) \) is the average measured ellipticity at the temperature maximum (\( T_{max} \)) (29). The \( m_{nd} \) and \( m_{d} \) are the average slopes for the first and last 10-15 °C temperature points respectively. The \( m_{nd} \) represents the temperature dependence of the ellipticity of the non-denatured (nd) DNA and \( m_{d} \) represents the temperature dependence of
ellipticity the denatured (d) DNA (29). This procedure corrected for the temperature
dependence of the upper and lower baselines.

The $\alpha$ expression in equation 5.4 was substituted into equation 5.5 and used to
analyze the thermal melt profiles of RHX and R*HX. The enthalpy and entropy terms
were obtained from the curve fit analysis and were then subsequently used to calculate
the melting point ($T_m$). The following general equation for $T_m$ can be applied to a non-
self-complementary system of any molecularity (28).

$$\frac{1}{T_m} = \frac{(n-1)R}{\Delta H} \ln C_T + \frac{[\Delta S-(n-1)R \ln(2n)]}{\Delta H}$$  \hspace{1cm} (Equation 5.6)

The $T_m$ expression for RHX and R*HX, where $n=3$, can be defined as:

$$T_m = \frac{\Delta H}{\Delta S+2R \ln \left(\frac{C_T}{6}\right)}.$$  \hspace{1cm} (Equation 5.7)

The Gibbs free energy ($\Delta G$), was calculated at 37 °C because it is physiologically
relevant (29).

$$\Delta G = \Delta H - T\Delta S = -RT \ln K_{eq}$$  \hspace{1cm} (Equation 5.8)

The melt profiles were individually fit to obtain $\Delta H$, $\Delta S$ and calculate $\Delta G_{37C}$ and $T_m$
values. The reported $\Delta H$, $\Delta S$, $\Delta G_{37C}$ and $T_m$ are average values based on a minimum
of 3 trials. The error for parameters was calculated as standard error of the mean.
Analysis of the RHX₄B Melt Profiles

Analysis of the RHX₄B melt profiles was done using Mathematica 6.0. For the 4-arm structure where n=4, the $K_{eq}(T)$ expression for the association structure can be expressed as (28)

$$K_{eq}(T) = \frac{[\text{RHX}_4\text{B}]}{[\text{R}][\text{H}][\text{X}_4][\text{B}]} = \frac{\alpha}{\left(\frac{C_T}{n}\right)^{n-1}(1-\alpha)^n} = \frac{64\alpha}{(C_T)^4(1-\alpha)^4} \, .$$

(Equation 5.9)

Mathematica 6.0 was used to solve equation 5.9 in terms of $\alpha$, which yielded only one root which was real in the range of the experimental data. Equation 5.10 shows the $\alpha$ expression after the van’t Hoff substitution for $K_{eq}$. 

182
$$\alpha = 1 + \sqrt{2}$$

\[\frac{1}{2} \left( \left( \frac{2}{3} \right)^{1/3} \right) - \left( \left( \frac{2}{3} \right)^{2/3} \right) \sqrt{3} \left( \frac{4C_1^2 \exp \frac{-\Delta H}{RT} R}{C_1^2 \exp \frac{-\Delta S}{RT} R} \right) - \sqrt{3} \left( \frac{27C_1^3 \left( \exp \frac{-\Delta H}{RT} R \right) - 4C_1^3 \left( \exp \frac{-\Delta S}{RT} R \right)^3}{C_1^2 \exp \frac{-\Delta S}{RT} R} \right) + \sqrt{3} \left( \frac{9C_1^3 \left( \exp \frac{-\Delta H}{RT} R \right) - 4C_1^3 \left( \exp \frac{-\Delta S}{RT} R \right)^3}{C_1^2 \exp \frac{-\Delta S}{RT} R} \right)\]

$$(\text{Equation 5.10})$$
Equation 5.10 was substituted into equation 5.5 and used to analyze the RHX₄B melt profiles. Using the fit values for ΔH and ΔS, the Tₘ of the 4-arm was determined using equation 5.12 (28).

\[
T_m = \frac{\Delta H}{\Delta S + 3R\ln\left(\frac{C_T}{8}\right)} \quad \text{(Equation 5.11)}
\]

The ΔG₃ʔC and its error were calculated by the same method used for the 3-arm data (29).

Analysis of the RRₑ Melt Profiles

For the RRₑ duplex (n=2) the Kₑq(T) for the association of the structure can be defined as:

\[
K_{eq}(T) = \frac{[RR_e]}{[R][R]} = \frac{\alpha}{(\frac{C_T}{n})^{n-1} (1-\alpha)^n} = \frac{2\alpha}{(C_T)(1-\alpha)^2} \quad \text{(Equation 5.12)}
\]

Mathematica 6.0 was used to solve equation 5.12 in terms of α. The negative form of the root was real in the range of the experimental data. The van’t Hoff expression for Kₑq was then substituted into the expression for α.

\[
\alpha(T) = \frac{1+C_T\exp\left(-\frac{\Delta H}{RT}\right)\sqrt{1+2C_T\exp\left(-\frac{\Delta H}{RT}\right)}}{C_T\exp\left(-\frac{\Delta H}{RT}\right)} \quad \text{(Equation 5.13)}
\]
Equation 5.13 was substituted into equation 5.5 and used to analyze the RRe melt profiles. The enthalpy and entropy terms were obtained from the curve fit analysis and were then subsequently used to calculate the melting point ($T_m$). For RRe, the $T_m$ expression was defined as:

$$T_m = \frac{\Delta H}{\Delta S + 4R\ln \left( \frac{C_T}{4} \right)} \quad \text{(Equation 5.14)}$$

The $\Delta G_{37C}$ and its error were calculated by the same method used for the 3-arm data (29).

**Analysis of the RHX Hybridization Profiles**

The formation of RHX occurs through a two step mechanism under the conditions studied here. The first step, the formation of duplex, is fast where $A+B \rightarrow AB$. The second step is the rate determining and observable step where $AB+C \rightarrow ABC$. Hence, the observable overall rate of the three way formation is given by:

$$r = k[AB][C] \quad \text{(Equation 5.15)}$$

where $[AB] = [C]$. As a result of the equimolar ratio, the equation can be simplified using $[S]$ which represents the strands that comprise the final structure (30).

$$r = k[S]^2 \quad \text{(Equation 5.16)}$$
Integration of the rate law and separation of variables yields an expression which describes \([S]\) at any time, \(t\), where at \(t=0\), the \([S]_t\) is equal to the sum of \([AB]_0\) and \([C]_0\).

\[
[S]_t = \frac{[S]_0}{1 + kt[S]_0} \quad \text{(Equation 5.17)}
\]

The \([S]_t\) is then converted into \(\Delta \varepsilon_{270}(t)\), the change in ellipticity at 270 nm.

\[
\Delta \varepsilon_{270}(t) = f_1 \left( \frac{[S]_0}{1 + kt[S]_0} \right) + f_2 \left( [S]_{\text{total}} - \frac{[S]_0}{1 + kt[S]_0} \right) \quad \text{(Equation 5.18)}
\]

In equation 5.18, \(f_1\) is the signal contribution from the non-complexed strands and \(f_2\) is the signal contribution from the 3-way structure, \(t\) is time in seconds, and \(k\) is the rate constant in \(\text{M}^{-1} \text{s}^{-1}\). The kinetic data curve-fit was done using Mathematica 6.0.

5.3 Results

5.3.1 Macrosolute Effects on the Thermodynamics of Multi-branch Junctions

Formation of the Multi-branch Structures

The RHX and R*HX sequences were derived from previously described 3-arm junctions (Figure 5.1) used by Clegg and colleagues to study the effects of unpaired bases on the global structure of 3-arm DNA (24). The 3-arm RHX structure is comprised of three mutually complementary sequences (25). Each of the three arms of RHX contain 10 base pairs and the arms are equidistance from each other (24).
R*HX system is also a 3-arm structure with a total of 30 base pairs. However, R*HX is more Y-shaped with two arms approaching each other as a result of three unpaired adenine bases present at the branch point (Figure 5.1) (24,31). The RHX₄B is a 4-arm structure with 10 base pairs per arm whose sequence is related to the 3-arm junctions and was also derived from a previously described structure (17). The RRₐ duplex is equivalent to two arms and is formed from the R strand of the RHX structure and its complement, Rₓ.

The formation of the RHX, R*HX and RHX₄B structures was confirmed by EMSA analysis as shown in Figure 5.2. The multi-branch structures, especially RHX₄B, were sensitive to salt concentrations. In order to confirm the structures, EMSA gels were made and run in 1X TBE buffer which contained NaCl to yield an ionic strength equivalent to PBS. In addition to the aforementioned structures, the R, H and X₄ strands were also studied to confirm that in the absence of one strand no higher order structures form. Only one band was produced for each structure which indicates only one product is formed upon annealing. The mobility of the bands was as expected, the 4-arm structure exhibited the slowest mobility while the duplex traveled the farthest. The bands of the RHX and R*HX structures co-migrated. The RHX₄, formed the expected structure whose mobility was similar to that of the 20 base pair RRₓ duplex.
Figure 5.1 – The sequences and schematic representation of the multi-branch junctions and the RRc duplex. (A) The RRc duplex was derived from the R strand of RHX and its complement, the R strand. (B) The 3-arm RHX structure has arms which are equidistant. (C) The 3-arm R*HX structure is more Y-shaped as a result of 3 unpaired bases and (D) The 4-arm RHX4B structure.
Figure 5.2– Confirmation of the multi-branch structures via EMSA. The 8% polyacrylamide gel (80:1 crosslink) was visualized with EtBr staining. Due to the salt sensitivity of the structures, the gels were prepared and run in buffer with added NaCl to obtain the ionic strength equivalent of PBS. The samples in lanes 5-9 were annealed prior to loading. (1) 6.5 µM R strand (2) 6.5 µM H strand (3) 6.5 µM X₄ strand (4) 6.5 µM B strand (5) 6.5 µM RRc duplex (6) 6.5 µM R*HX (7) 6.5 µM RHX₄ (8) 6.5 µM RHX (9) 6.5 µM RHX₄B

Macrosolute Effects on the Thermodynamic Properties of RHX and R*HX

The thermal stability of RHX and R*HX were studied in a dilute environment consisting of buffer (PBS) and in a crowded environment created by the presence of a macrosolute (20% PEG 8K in PBS). The macrosolute selection was based on previously noted enhancement of duplex DNA in 20% PEG 8K (chapters 2 and 3) and on reports of stabilization of duplex DNA in PEG 8K (3-6,32-34). The thermal denaturation of all the structures was monitored via change in ellipticity at 270 nm. In Figure 5.3a, representative wavelength scans for RHX in PBS are shown. The scans were taken before and after a melt at 25 and 75 °C respectively. The monitoring wavelength was selected because at 270 nm, the greatest change in signal upon complete denaturation of the structures was noted. The wavelength scans indicate that
RHX is in a B-form conformation as shown by the characteristic maximum and minimum CD peaks at 245 and 280 nm respectively (35).

The temperature at which half of the structure has denatured and is present in the single strand form is the melt temperature ($T_m$). The representative melts in Figure 5.3b depict the two-state melting and macrosolute induced stabilization of RHX. In PBS, the $T_m$ of RHX was 49.3 °C and increased to 56.8 °C in the presence of the macrosolute (Table 5.2). The difference in the $T_m$ of RHX in the crowded compared to the dilute environment ($\Delta T_m$) was 7.5 °C and demonstrated the ability of the macrosolutes to substantially stabilize a multi-arm structure. The Gibbs free energy ($\Delta G_{37C}$) is a thermodynamic parameter that can be used as a standard for comparing various structures that melt at different temperatures. The difference in stability resulting from crowding ($\Delta \Delta G_{37C}$) was -3.9 kcal mol$^{-1}$, a significant stabilization due to the presence of the macrosolute.

To investigate what role shape may play in the stabilization, R*HX was studied. In contrast to the equidistance arms of RHX, the arms in the R*HX structure are closer together and more Y-shaped (24). Figure 5.3c shows wavelength scans of R*HX in PBS taken before and after a melt at 25 and 95 °C respectively. The characteristic CD peak minimum and maximum are equivalent to those of the RHX structure indicating no loss of B-form structure as a result of the unpaired bases. Additionally, the large change in signal at 270 nm, indicative of the loss of structure, is also shown in the pre versus post-melting wavelength scans.
Figure 5.3 – Representative wavelength scans and melt profiles for the 3-arm structures and the RR<sub>c</sub> duplex. Scans were carried out at the start (blue) and end (red) of melts. Representative melts in PBS (open circle) and 20% PEG 8K (open square) were curve fit as depicted by the black and blue lines respectively. (A) RHX scans at 25 and 75 °C (B) RHX melts (C) R*HX scans at 25 and 75 °C (D) R*HX melts (E) RR<sub>c</sub> duplex scans at 55 and 95 °C (F) RR<sub>c</sub> duplex melts.
Table 5.2 – The thermodynamic parameters of hybridization for the RR<sub>c</sub> duplex, RHX, R*HX and RHX₄B structures. All values were determined from the thermal denaturation studies of the structures in the presence of PBS and 20% PEG 8K. The values shown are based on the average of three or more individually fit trials. Reported errors were calculated as the standard error of the mean. The Δ<sub>Tₕ</sub> and ΔΔ<sub>G₃₇C</sub> errors were tabulated using propagation of error.

The melt profiles of R*HX were conducted in the same manner as the melts of RHX, at 270 nm in PBS and in the presence of 20% PEG 8K (Figure 5.3d). The addition of an unpaired base in duplex DNA can substantially destabilize the structure (36). However, the additional three adenine bases in R*HX do not effect its stability as demonstrated by the identical T<sub>ₕ</sub>’s of RHX and R*HX (Table 5.2). These results are in agreement with previous reports, by Clegg and colleagues, that RHX and R*HX have

<table>
<thead>
<tr>
<th></th>
<th>RR&lt;sub&gt;c&lt;/sub&gt; Duplex</th>
<th>RHX</th>
<th>R*HX</th>
<th>RHX₄B</th>
</tr>
</thead>
<tbody>
<tr>
<td>ΔH (kcal mol&lt;sup&gt;-1&lt;/sup&gt;)</td>
<td>-151.4 ± 0.2</td>
<td>-147.4 ± 1.3</td>
<td>-148.1 ± 0.1</td>
<td>-168.3 ± 0.3</td>
</tr>
<tr>
<td>ΔS (cal mol&lt;sup&gt;-1&lt;/sup&gt;K&lt;sup&gt;-1&lt;/sup&gt;)</td>
<td>-402.8 ± 0.9</td>
<td>-403.2 ± 3.8</td>
<td>-405.3 ± 0.6</td>
<td>-455.8 ± 0.8</td>
</tr>
<tr>
<td>Tₕ (°C)</td>
<td>79.2 ± 0.4</td>
<td>49.3 ± 0.3</td>
<td>49.3 ± 0.3</td>
<td>38.5 ± 0.3</td>
</tr>
<tr>
<td>ΔG&lt;sub&gt;₃₇C&lt;/sub&gt; (kcal mol&lt;sup&gt;-1&lt;/sup&gt;)</td>
<td>-26.5 ± 0.1</td>
<td>-22.4 ± 0.2</td>
<td>-22.4 ± 0.1</td>
<td>-26.9 ± 0.2</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>RR&lt;sub&gt;c&lt;/sub&gt; Duplex</th>
<th>RHX</th>
<th>R*HX</th>
<th>RHX₄B</th>
</tr>
</thead>
<tbody>
<tr>
<td>ΔH (kcal mol&lt;sup&gt;-1&lt;/sup&gt;)</td>
<td>-159.3 ± 0.6</td>
<td>-158.3 ± 1.1</td>
<td>-157.6 ± 0.8</td>
<td>-190.8 ± 0.5</td>
</tr>
<tr>
<td>ΔS (cal mol&lt;sup&gt;-1&lt;/sup&gt;K&lt;sup&gt;-1&lt;/sup&gt;)</td>
<td>-420.6 ± 1.8</td>
<td>-425.6 ± 3.6</td>
<td>-422.1 ± 2.0</td>
<td>-505.2 ± 1.8</td>
</tr>
<tr>
<td>Tₕ (°C)</td>
<td>82.7 ± 0.1</td>
<td>56.8 ± 0.3</td>
<td>56.8 ± 0.4</td>
<td>50.7 ± 0.3</td>
</tr>
<tr>
<td>ΔG&lt;sub&gt;₃₇C&lt;/sub&gt; (kcal mol&lt;sup&gt;-1&lt;/sup&gt;)</td>
<td>-28.8 ± 0.1</td>
<td>-26.2 ± 0.1</td>
<td>-26.5 ± 0.2</td>
<td>-34.2 ± 0.2</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>ΔTₕ (°C)</th>
<th>ΔΔG&lt;sub&gt;₃₇C&lt;/sub&gt; (kcal mol&lt;sup&gt;-1&lt;/sup&gt;)</th>
</tr>
</thead>
<tbody>
<tr>
<td>PBS</td>
<td>3.5 ± 0.4</td>
<td>-2.3 ± 0.1</td>
</tr>
<tr>
<td>20% PEG 8K</td>
<td>7.5 ± 0.4</td>
<td>-3.8 ± 0.2</td>
</tr>
<tr>
<td>∆ (PEG 8K – PBS)</td>
<td>7.5 ± 0.5</td>
<td>-4.1 ± 0.2</td>
</tr>
<tr>
<td></td>
<td>12.2 ± 0.4</td>
<td>-7.3 ± 0.3</td>
</tr>
</tbody>
</table>
The RHX and R*HX structures also had similar $\Delta G_{37C}$ values of -22.4 kcal mol$^{-1}$. The macrosolute induced increase in $T_m$ for R*HX was 7.5 °C, identical to the $\Delta T_m$ of RHX in macrosolute, and the resulting $\Delta \Delta G_{37C}$ for R*HX is -4.1 kcal mol$^{-1}$, a value similar to the stability increase of RHX in the crowded solution.

The effects of crowding on duplex DNA is typically 3-5 °C in the presence of macrosolutes (4,5). However, the effects of crowding noted here for the 3-arm structures are greater than that previously reported for duplex DNA. Crowding effects may be sequence dependent, hence to investigate how molecularity and thermodynamic crowding effects are linked, the RR$_c$ duplex was created. The RR$_c$ duplex is related in sequence to the 3-arm structure, as RR$_c$ is comprised of the R strand from RHX and its complement.

Wavelength scans demonstrate that RR$_c$ is a B-form duplex (Figure 5.3e). The RR$_c$ melting was also monitored by the changes in ellipticity at 270 nm in PBS and 20% PEG 8K. The 20 base pair duplex had a $T_m$ of 79.2 °C in PBS. The crowding effect on the duplex was not as great as that of the 3-arm structures. In the presence of macrosolute, the $T_m$ for the RR$_c$ duplex increased by 3.6 °C (Figure 5.3f and Table 5.2). Comparison of the $\Delta \Delta G_{37C}$ confirms the greater stabilization for the RHX 3-arm structure ($\Delta \Delta G_{37C}$ was -3.9 kcal mol$^{-1}$) over the duplex ($\Delta \Delta G_{37C}$ was -2.3 kcal mol$^{-1}$).
The Thermodynamic Effects of Crowding on RHX₄B

Crowded conditions resulted in stabilization of the 3-arm junctions and the RRₜ duplex. The 4-arm junction, RHX₄B, is closely related to the 3-arm junction, RHX. The fourth arm on RHX₄B was created by the addition of the B strand and modification of the X strand (Figure 5.1). Since crowding effects may be sequence dependent, studying related sequences such as RRₜ, RHX and RHX₄B reduces effects resulting from varying sequences. Thus, the relationship between molecularity and crowding effects could be better isolated.

The arms of RHX₄B were demonstrated to be in a B-form conformation as indicated by the characteristic CD peak maximum and minimum (Figure 5.4a). For direct comparison purposes RHX₄B thermal denaturation was monitored by changes in ellipticity at 270 nm in PBS and in the presence of 20% PEG 8K (Figure 5.4b). The Tₘ of RHX₄B was 38.5 in PBS. Thus, at similar total strand concentrations, RHX₄B has a lower Tₘ than the other structures due to its higher molecularity. However, the crowding mediated stabilization of RHX₄B was greater than that of the 3-arm junctions and the RRₜ duplex. In the presence of PEG, the Tₘ of RHX₄B was increased by 12.2 °C (Table 5.2). The RHX₄B structure was also stabilized by 7.5 kcal mol⁻¹ in the presence of macrosolute as indicated by the ΔΔG₃⁰°C (Table 5.2).
Figure 5.4 – Representative wavelength scans and melt profiles of RHX$_4$B. (A) RHX scans at 15 and 75 °C as denoted by the blue and red circles respectively. (B) Representative melt profiles of RHX in PBS (open circle) and in the presence of 20% PEG 8K (open square). The PBS and PEG melt profiles were curve fit as depicted by the black and blue lines respectively.

5.3.2 Macrosolute Effects on the Kinetics of RHX Formation

In addition to thermodynamic consequences, crowding can also effect the kinetics of structure formation by increasing the rate of the forward reaction. The addition of cosolutes to a system results in an increase in the viscosity and would be expected to decrease the rate of association (37). However, an increase in the chemical activity of reactants in the presence of cosolutes may oppose viscosity effects (38). The kinetics of RHX formation was studied in PBS (dilute environment) and in the presence of 20% PEG 8K (crowded environment) in order to investigate how a multi-component structure’s rate of formation would be effected by crowding. The kinetics of RHX formation have not been previously quantified, but were qualitatively...
reported as being much slower than the kinetics of duplex formation (24). The inherently slow kinetics of multi-component structure and the low concentrations may be why it was not previously possible to observe the kinetics of formation. Here we report the kinetics of RHX formation which was monitored via CD using μM concentrations of the strands.

The assembly of RHX appears to occur through a two step mechanism. The overall mechanism involves formation of a duplex with single strand overhangs, followed by hybridization of a third strand to the overhanging regions. The first step, the formation of a duplex, is fast and unobservable in our system at the μM concentrations employed (39,40). The second step is the formation of the 3-arm which is the slow and observable step. As a result the data was fit to a second order rate law which represents the rate of the slower second step of the mechanism.

The RHX kinetic data was obtained by monitoring the change in ellipticity at 270 nm, upon mixing of the separate strands in an equimolar ratio. In PBS the rate constant for formation of RHX was $63.5 \pm 1.5 \, \text{M}^{-1}\text{s}^{-1}$ (Figure 5.5a). The kinetics of RHX formation in the dilute environment (PBS) was orders of magnitude slower than the typical rate of duplex formation which is $10^5$ to $10^7 \, \text{M}^{-1}\text{s}^{-1}$ (39). The slow rate of structure formation highlights the increased barrier for hybridization of the third strand in RHX. In the presence of 20% PEG 8K, the rate constant increased by almost 10-fold to $602 \pm 10.6 \, \text{M}^{-1}\text{s}^{-1}$ (Figure 5.5a).
Figure 5.5 – Macrosolute effects on the kinetics of RHX formation. (A) The average rate of RHX formation in PBS (grey) and in the presence of 20% PEG 8K (blue). (B) The addition of all strands simultaneously (blue) and the “2+1” addition of the strands (red) yields similar plots supporting the proposed mechanism of RHX formation. (C) The average rate of 6.5 µM RHX (grey) and 9.19 µM RHX (orange) kinetics of hybridization in PBS. The increase in concentration resulted in the expected increase in the rate of hybridization while the rate constant remained unchanged, supporting second order nature of the observed hybridization. The solid lines represent the second order fit. All data shown is based on the average of 3 or more trials.
Support for the proposed mechanism of RHX formation was obtained from additional CD studies in which the strands were mixed in a “2+1” manner. Two of the strands, R and X were incubated for an hour after which the H-strand was added and data collection was initiated. The “2+1” kinetics of RHX formation were monitored by the change in ellipticity at 270 nm in 20% PEG 8K. The data was treated and curve fit identically to the RHX kinetics. The value obtained for the “2+1” rate constant in 20% PEG 8K (601 ± 10.4 M⁻¹s⁻¹) was in agreement with the rate constants obtained previously where the strands were simultaneously mixed (602 ± 10.6 M⁻¹s⁻¹) (Figure 5.5b). Furthermore, when the concentration of RHX strands in PBS was increased from 6.5 µM to 9.19 µM, the average rate constant of the 9.19µM reaction (70 ± 1.5 M⁻¹s⁻¹) remained nearly the same as the 6.5 µM (63.5 ± 1.5 M⁻¹s⁻¹). These results confirm the second order behavior of the RHX hybridization reaction (Figure 5.5c).

EMSA analysis confirmed that the product of the RHX kinetics was the same as that of the thermodynamically studied (annealed) 3-arm structure. The migration of the band for the annealed RHX sample matched that of the RHX kinetic sample. The migration of both RHX samples was significantly less than the single strand and duplex bands (Figure 5.6). EMSA was also used to provide further support for the proposed mechanism. EMSA analysis confirmed the rapid rate of duplex formation at the same concentrations employed for the kinetic studies via CD. The three possible duplexes formed from the RHX strands were mixed and immediately loaded onto a running gel.
Each of the three samples produced bands that co-migrated with the annealed duplex samples, and all duplex migrated slower than the single strand components (Figure 5.6).

**Figure 5.6 – Confirmation of the RHX kinetic product.** All samples were 6.5 μM per strand in PBS. The red labels denote annealed samples and the blue labels denote samples which were incubated. The 8% polyacrylamide gel was visualized with EtBr staining. (1) R strand (2) H strand (3) X strand (4) HX duplex (incubation time > 1 minute prior to loading) (5) annealed HX duplex (6) RX duplex (incubation time > 1 minute prior to loading) (7) annealed RX duplex (8) RH duplex (incubation time > 1 minute prior to loading) (9) annealed RH duplex (10) XH+R (1 hour incubation for X+H followed by a 1 hour incubation of XH+R) (11) RX+H (1 hour incubation for X+R followed by a 1 hour incubation of RX+H) (12) RH+X (1 hour incubation for R+H followed by a 1 hour incubation of RH+X) (13) RHX (2 hour incubation) (14) annealed RHX. The gel demonstrates that the product of RHX incubation is identical to the annealed product. Furthermore, the gel confirms the rapid formation of duplex.
5.4 Discussion

5.4.1 Macrosolute Induced Stability Correlates with Molecularity

The RR<sub>c</sub> duplex was used to study the effects of crowding on system where n=2. The results shown here for the RR<sub>c</sub> duplex, a 3.5 °C increase in stability in the presence of macrosolute, is in agreement with expectations based on reported values (4). While the effects of crowding on triplex DNA have been studied, the effects of crowding on a canonical branched structure such as RHX and R*HX (n=3) have not been reported previously. Our investigation of the n=3 structures allowed for comparison of molecularity effects on crowding mediated stabilization of nucleic acids. The observed increase in stability for both RHX and R*HX, as indicated by the ΔΔG<sub>37</sub>, was almost 2-fold greater than the stability noted for the RR<sub>c</sub> duplex in the presence of macrosolute. The size and shape of cosolutes are important since increasing size and globular-like shapes can result in increased crowding effects on nucleic acid systems (3,5). The R*HX structure was Y-shaped as a result of unpaired bases which did not destabilize the structure. Here, the difference in shape between the RHX and R*HX species did not result in a difference in observed crowding effects.

In the presence of macrosolute, RHX<sub>4</sub>B demonstrated a 12.2 °C increase in T<sub>m</sub>. Previous reports of such large increases in T<sub>m</sub> have been documented for triplex DNA by Minsky and Chaires and Spink (3,5). In Minsky’s study of 20 base pair triplex forming oligonucleotides in the presence of 20% PEG 8K, he reports up to 15 °C increase in T<sub>m</sub> of a crowded versus a dilute environment (3,4). Chaires and Spink show
that in a crowded environment a triplex could exhibit an increase in T_m of up to 20°C in the presence of macrosolutes (5). However, here we present a stabilization of a multi-arm junction, with canonical B-form arms, that exhibits stabilization similar to that of a triplex. Unlike a triplex, which has very strict base pairing and pH requirements, the 4-arm structure formation does not rely on Hoogsteen base pairing and is less restricted.

The observation of increasing thermodynamic crowding effects with increasing molecularity is relevant because of the related sequences. The relationship between crowding effects and molecularity is not strictly dependent on the number of base pairs. It has been suggested that as the number of base pairs in a duplex increase, the observed crowding mediated stabilization increases as well. However, Spink and Chaires studied a 20 base pair duplex in 20% PEG 8K which exhibited a 5°C increase in T_m (5). In the same study, a 200 base pair duplex showed a 6°C increase in T_m under the same conditions. Therefore, the increase noted here for the n=3 and n=4 duplexes could not be solely due to the change in the number of base pairs.

Crowding will favor the compact form of a structure. The co-volume of a duplex as compared to the co-volumes of the single strands is reduced. The reduction in co-volume results in a greater available volume for the cosolutes which can lead to stabilization of the duplex. The reduction of the co-volume for a multi-branched structure (compared to the individual strands) correlates with increasing molecularity. We postulate that the crowding effects observed for the multi-branch junctions are
mostly due to the excluded volume effects of the macrosolutes as a result of the increasing reduction in co-volumes between the 3 and 4-arm structures.

**Kirkwood-Buff Analysis of the Thermodynamic Data**

The Kirkwood-Buff theory of solutions (41) has been used by Shimizu and co-workers to interpret the crowding mediated effects observed for the multi-branch structures (42-45). In the Kirkwood-Buff interpretation, the system is described as consisting of water (i=1), a macromolecule (i=2) and cosolute (i=3) (section 1.8). Using these components, the preferential hydration parameter ($\nu_{21}$) can be described as the effect that water, $\mu_1$ (in terms of chemical potential), has on a macromolecule, at a constant pressure (P) temperature (T) and molality of the macromolecule ($m_2$) (41,43,46).

$$\nu_{21} = \left( \frac{\partial \mu_2}{\partial \mu_1} \right)_{P,T,m_2}$$ (Equation 5.19)

When the macromolecule undergoes a reaction or change in conformation, the change in the preferential hydration parameter, $\Delta \nu_{21}$, can be expressed as the change in the equilibrium constant, $K_{eq}$, or as the change in the number of cosolutes ($\Delta N_{23}$), and water molecules ($\Delta N_{23}$) distributed around the macromolecule in terms of molality (equation 5.21) (42).

$$\Delta \nu_{21} = RT \left( \frac{\partial \ln K}{\partial \mu_1} \right)_{T,P,n_2} = \Delta N_{21} - \frac{m_1}{m_3} \Delta N_{23}$$ (Equation 5.20)
\[ \Delta N_{21} = n_i N_A \Delta G_{21} = n_i N_A \left( G_{21}^B - G_{21}^A \right) \]  
\[ \text{(Equation 5.21)} \]

\[ G_{2i}^\sigma = \int d \bar{r} [g_{2i}^\sigma(\bar{r}) - 1] \]  
\[ \text{(Equation 5.22)} \]

The \( \Delta N_{21} \) and \( \Delta N_{23} \) are directly correlated to the radial distribution of the water (\( G_{21} \)) and cosolutes (\( G_{23} \)) around the macromolecule respectively (equations 5.21 and 5.22).

In equation 5.22, \( N_A \) is Avogadro’s number, \( g_{2i}^\sigma(\bar{r}) \) is the function of correlation between the macromolecule in state \( \sigma \) and species \( i \), and is written as a function of \( \bar{r} \), the radial distance from the macromolecule (42).

According to Shimizu, macrosolute effects are based on cosolute exclusion, while the effects of osmolytes are driven by changes in hydration (47). Kirkwood-Buff is a robust theory for interpreting cosolute effects because it can be used to interpret the effects of both macrosolute and osmolytes (45). Other theories such as osmotic stress and preferential hydration are only reliable for osmolytes and interpret cosolute effects as purely hydration based (48). Additionally, osmotic stress analysis has been shown to overestimate the number of water molecules involved in a reaction (43).

In order to estimate the \( \Delta \nu_{21} \), published data was used which reported changes in water activity as a result of the presence of 20% PEG 8K (34). Since the water activity values were obtained from experiments whose buffer was different from those studied here, the \( \Delta \nu_{21} \) values calculated here are only general approximations (Table 5.3). The conclusions drawn are strictly qualitative but valuable because they provide insight into the mechanism of cosolute induced effects. The \( K_{eq} \) (37 °C) values were tabulated.
using the calculated $\Delta G_{37\degree C}$ values previously calculated from the thermodynamic studies of the multi-branch junctions (Table 5.2). The $\Delta K_{eq}$ were obtained by comparing the $K_{eq}$ values in crowded versus dilute solutions and then used to evaluate $\Delta v_{21}$.

<table>
<thead>
<tr>
<th></th>
<th>$\Delta T_m$ ($\degree C$)</th>
<th>$\Delta \ln K_{eq}$ (mol$^{-1}$)</th>
<th>$\Delta v_{21}$ (mol mol$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>RR$_c$</td>
<td>3.6</td>
<td>0.9</td>
<td>Negative</td>
</tr>
<tr>
<td>RHX</td>
<td>7.5</td>
<td>1.4</td>
<td>Negative</td>
</tr>
<tr>
<td>R$^*$HX</td>
<td>7.5</td>
<td>1.6</td>
<td>Negative</td>
</tr>
<tr>
<td>RHX$_d$B</td>
<td>12.2</td>
<td>2.8</td>
<td>Negative</td>
</tr>
</tbody>
</table>

**Table 5.3 – The change in preferential hydration upon addition of cosolutes.** The change in preferential hydration parameter was estimated using equation 5.20. Previously published values for change in water activity and the thermodynamic parameters from Table 5.2 were used to estimate $\Delta v_{21}$.

For all of the structures studied here the $\Delta v_{21}$ was negative. As seen in equation 5.20, $\Delta v_{21}$ is based on the change in the number of water ($\Delta N_{21}$) and the change in the number of cosolutes ($\Delta N_{23}$) distributed around the DNA. Hybridization of DNA involves the uptake of water while the disruption of a duplex results in the release of approximately 4 water per base pair (5,32). The thermodynamic parameters reported in Table 5.1 are values for hybridization of the structures, hence, for the formation of the structures studied here $\Delta N_{21}$ is positive. Therefore, the negative $\Delta v_{21}$ values indicate that the $\Delta N_{23}$ must be positive and larger than $\Delta N_{21}$. The positive $\Delta N_{23}$ is an indicator
that the PEG distribution is greater around the final structures than around the single strands (Figure 5.7).

Although it has been previously stated that crowding effects are based solely on changes in hydration (48), the results shown here indicate that that the consequences of cosolute distribution are greater than that of water. However, attributing the enhancement solely to $\Delta \nu_{21}$ would overestimate $\Delta N_{21}$ (45). The presence of macrosolutes does in fact alter water activity, but this effect is significantly outweighed by the changes in the macrosolute exclusion. Hence we conclude the stabilization effects arise mostly from excluded volume effects with minor contributions from changes in the distribution of water.

**Figure 5.7 - A schematic representation of the distribution of macrosolutes and waters upon duplex formation and dissociation.** The larger circles represent a unit of volume composed of water (i=1, blue spheres), DNA (i=2) and 20% macrosolute (i=3, red spheres). Bulk solvent properties are observed outside of the dashed circle. The inner circle represents the immediate vicinity surround DNA, the local domain. In environments crowded with macrosolutes, upon hybridization the $\Delta N_{21}$ is positive, but smaller than the $\Delta N_{23}$. Hence, the macrosolute distribution around the vicinity of the duplex is greater than around the single strands. For clarity, only a representative number of water molecules are shown.
5.4.2 Macrosolute Presence Enhances the Rate Constant of RHX Formation

The kinetics of RHX formation was studied in crowded and dilute environments. We conclude that the formation of RHX likely occurs through two steps, resulting in the final structure. In our system, the observable portion of the mechanism could be treated like that of bimolecular system and thus the formation of RHX data was fit with a second order rate expression. The overall rate of formation for RHX was two orders of magnitude slower than that for a duplex of similar base pairs. The slow rate of formation highlights the complexity of the RHX structure and high activation energy. In the presence of macrosolute, the enhancement of RHX rate of formation was 10-fold greater than in the dilute environment. We have found that the effects of crowding on a linear 16-mer duplex structure can result in up to a 3-fold increase in rate in 20% PEG 8K (chapter 2). However, the effect of macrosolutes on hybridization of complex structures (stem-loop and target duplexes) increases of up to 12-fold have been observed (chapter 4). Comparison of the macrosolute enhancement with other complex systems reveals that the enhancement for RHX is similar. Hence, macrosolute effects on the kinetics of RHX formation may scale with the activation energy of the system.
5.5 Conclusions

The evidence presented here indicates a crowding trend which relates increasing molecularity to increased crowding mediated thermodynamic effects (8). Additionally, it appears that the observable effects of crowding on kinetics scales with the activation energy of the structures. We conclude that the thermodynamic crowding mediated trend presented here is not due to the increasing base pairs but rather increasing molecularity. Furthermore, we conclude that the crowding mediated effects are mostly due to excluded volume effects caused by the macrosolutes. The largest thermal stability enhancement noted here for the RHX₄B system is of similar magnitude to the stabilization previously reported for triplex DNA stabilization in the presence of macrosolutes. However, although triplex DNA is more compact than duplex DNA, the multi-branch structures presented here contain canonical B-form arms. The work presented here demonstrates that macromolecular crowding is a factor that influences the thermodynamic and kinetic properties of DNA. In addition, the work presented here can be applied to nanostructures, biosensors, nanomotors or other higher complexity structures.
5.6 Technological and Physiological Implications of Macromolecular Crowding

Technological Implications

5.6.1 Nanostructure Design

DNA is genetic material but it is possible to utilize it as a generic material in the formation of nanostructures (22). Recently, nanostructures have become of great interest due to their potential applications as scaffolds, biosensors, and nanomotors (22,49,50). The simplest nanomotor is commonly referred to as a “nanoswitch” which has two possible conformations (23). The state of the nanoswitch is controlled by changes in salt composition or the by the addition of specific DNA fuel strands (51).

One of the simplest and most common nanostructures is the conversion of a duplex to a triplex (52). The conversion takes place when an appropriate sequence of single stranded DNA is added to a duplex at pH 5.0. Another example of a nanodevice includes a closed loop DNA duplex which was attached to opposite arms of a Holliday junction (53). Upon addition of ethidium bromide, a conformational change of the DNA loop was induced and the helix was partially unwound. The change in conformation of the loop resulted in branch exchange of the Holliday junction. In addition to simple switches of conformation, nanodevices can “walk” on the surfaces of macromolecules. A bipedal walking nanodevice was created by Shin and Pierce using a double-stranded duplex as a track and a duplex with overhanging regions as the
bipedal motor (54). The walking motion was induced by adding “attaching” and
“detaching” fuel strands in alternating order.

In the previous examples, a change in the experimental conditions was required
to produce a conformational change in a nanoswitch, or the addition of a fuel strand was
necessary to induce motion of a nanodevice. However, it may be possible to exploit
the crowding effects, observed in this chapter on the multi-branch systems, to create
switchable systems. Work by Sugimoto provides an example of a nanoswitch that was
controlled by a crowded environment (55). Sugimoto reports the conversion of an anti-
parallel G-quadruplex to parallel G-quadruplex when the structure was placed in a
crowded environment. Here we have demonstrated more dramatic changes in structure
(assembly of 3 or 4-arm structures). DNA structures with high molecularity (>3) are
very sensitive to crowding conditions and thus it may be possible to create a
nanostructure that does not spontaneously form in a dilute environment, but does self-
assemble in the presence of macrosolutes.

Physiological Implications

5.6.2 Single Nucleotide Polymorphism Detection

Single nucleotide polymorphisms (SNPs) are variations that occur in the
sequence of DNA where a single nucleotide is replaced by one of the other possible
three (56). SNP classification is only given to a variation when the single nucleotide in
the genome is different in at least 1% of a species (57). A total of 1.4 million SNPs were reported in the human genome project findings, with a SNP occurring every 1000-2000 bases on average (58). A majority of SNPs are in non-coding regions, and some may alter an individual’s response to pathogens or even to drug therapy (59). SNPs in coding regions have been implicated in various diseases such as cancer (60).

SNPs can be detected in both sequence nonspecific or sequence specific manners. Sequence nonspecific detection is a powerful tool for SNP discovery (61). The sequence non-specific detection exploits the differences between mismatched DNA duplexes and single-stranded DNA via electrophoresis, cleavage or chromatography (62). However, sequence specific detection is useful and preferred for genotyping. Allele specific oligonucleotide hybridization assays are an example of a sequence specific method for SNP detection (61). If a SNP is present on the target a single mismatch will be formed between the target and probe. The probe is designed such that is will preferentially hybridize to its fully complementary targets. Hybridization between a probe and SNP target is disfavored as a result of the destabilization caused by the single mismatch.

The challenge that exists in allele specific oligonucleotide hybridization SNP genotyping is that probe design can be difficult. However, as demonstrated in chapter 2, the energy difference between a WT and single mismatched base containing duplex can be widened. A WT duplex was differentially stabilized over two single mismatch
duplexes by \(-1.7\ \text{kcal mol}^{-1}\) in the presence of a macrosolute. Thus, it may be possible to increase the specificity of hybridization with the use of crowded conditions.

Allele-specific probes can be immobilized onto a solid support in a SNP array (63). The solid support can be a latex bead, a silicon chip, or a well of a 96-well plate. In this method, hybridization of labeled target DNA is the manner of detection (62). An advantage to the SNP array is that a variety of targets can be incorporated on one support (64). It is possible to enhance the specificity of SNP arrays with macrosolutes, however, additional modifications to the method would be required. Our experiments demonstrated an increase in hybridization rate (a limiting factor at nM concentrations of DNA) such that macrosolutes could theoretically decrease the time required for hybridization. We previously conducted enzyme linked immunosorbent assays (ELISAs) in microtiter plates in the presence of macrosolute (data not shown). In this heterogeneous system, the presence of macrosolute caused mass transport to the surface to become the limiting factor. Hence, SNP arrays in a crowded environment could still show advantages but would require active recirculation of the buffer/crowded solution in order to overcome mass transport obstacles.

5.6.3 MicroRNA Implications

MicroRNAs (miRNAs) are \(~22\) nucleotide non-coding RNA single strands that regulate gene expression and may play a role in the pathogenesis of cancer (65). The
miRNAs can bind to the regions on the 3’ end of untranslated mRNAs (66). The first step in miRNA processing occurs when they are transcribed by RNA polymerase II as primary miRNAs (pri-miRNAs), stem-loops with 5’ caps and 3’ poly(A) tails (67). The next step involves Drosha, a double-stranded RNA-specific ribonuclease. Drosha digests the pri-miRNA in the nucleus to form precursor miRNAs (pre-miRNAs). After Drosha processing, either the 5’ or 3’ end of the mature miRNA is generated (68). The pre-miRNA is exported from the nucleus to the cytoplasm by Exportin-5 (69). Then the pre-miRNA is cleaved by Dicer (an RNase III enzyme) which results in small double-stranded RNA with 3’ overhangs of 1-4 nucleotides (69). The RNA duplexes contain mismatched bases and bulged bases. The double-stranded RNA contains the mature miRNA strand which upon separation of the strands, is associated with RNAi-induced silencing complex (RISC), and becomes active miRNA (67). The remaining strand of the RNA duplex is degraded by RISC (70).

As demonstrated in chapters 2 and 3, the presence of macrosolute generally stabilized duplexes by an average of 1-2 kcal mol$^{-1}$. The miRNA structures are ~22 nucleotides in length and are expected to have an average $\Delta G_{37C}$ of about -10-12 kcal mol$^{-1}$ (71). This corresponds to a dissociation constant close to 10 nM, therefore, in the concentration ranges likely present in cells, an additional 1-2 kcal mol$^{-1}$ resulting from crowding mediated effects may be significant in terms of binding. In addition, the effects of macrosolutes combined with mismatches and bulge bases can result in the
differential stabilization of one structure over another. Bulges and mismatches are a hallmark of miRNA and they will likely be sensitive to these differences.

The kinetics of duplex formation were also demonstrated to be effected by the presence of macrosolutes (chapters 2-5). Bulge and mismatch containing duplexes, stem-loop duplexes, and a 3-arm structure, exhibited rate of hybridization enhancement in the presence of macrosolute. This observation indicated that unlike thermodynamic effects, the kinetic effects are not particularly sequence dependent. Thus, in addition to the thermodynamic enhancement of the miRNA structures, the kinetics of the miRNA association to an mRNA target, and the association of the miRNA with the RISC complex is expected to be faster in vivo.

5.6.4 DNA Replication – Okazaki Fragments

In DNA replication, double stranded DNA is unwound and a replication fork is formed when the two strands are split (72). The two strands are called leading and lagging strands because of their directionality. The leading strand goes from 5’ to 3’ and DNA can be synthesized continuously (73). However, the lagging strand goes from 3’ to 5’ and DNA is synthesized in short fragments called Okazaki fragments (72). The fragments are typically 100-200 nucleotides in length and are very stable duplexes. As demonstrated by Chaires and Spink, the expected change in free energy for a duplex
of this length would be approximately 2 kcal mol\(^{-1}\) (32,74). Hence, in comparison to their inherent stability, crowding will not have a significant stabilizing effect.

5.6.5 – Resolution of Holliday Junctions

Holliday Junctions are formed when strand exchange occurs between two homologous duplexes (75). Holliday junctions are key intermediates in recombination processes. The interconnected structure can also be formed during replication if fork reversal occurs. Electrophoresis and FRET experiments were the first to show the formation of the Holliday junction (17,27). However, in the last ten years, the crystal structures of DNA Holliday junctions have been reported (76). Junction resolving enzymes can resolve a Holliday junction and create two duplexes by introducing paired nicks close the branch point of the structure (17,77).

The RHX\(_4\)B structure (chapter 5) was a 4-arm structure formed from four individual strands. The stabilization exhibited by the RHX\(_4\)B structure was about -7 kcal mol\(^{-1}\) in the presence of macrosolute. In the same conditions, the RR\(_c\) duplex was stabilized by about -2 kcal mol\(^{-1}\). The differential energy due to crowding between 4 strands and two duplexes is approximately twice that, or -4 kcal mol\(^{-1}\). We can estimate that the differential energy between the RHX\(_4\)B and two duplexes in a crowded environment would be approximately -3 kcal mol\(^{-1}\). Thus, in vivo, we would predict that the 4-arm structure would be stabilized by crowding over the two duplexes. Resolution
of Holliday junctions by junction-resolving enzymes, however, is still energetically favored since the combined free energy of two duplexes would be greater than the 4-arm structure. We conclude that the kinetics of branch migration of a Holliday junction would not be affected by the crowded physiological environment since there is no change in energy or covolume of the structure.

In vivo 4-arm structures can also be formed from two strands, also known as cruciform structures (73). Based on the crowding effect observed for duplexes (chapters 2-6), we conclude that in vivo, the cruciform would not be stabilized to the same extent as the Holliday junction due to its molecularity. Although in vivo cruciform’s can contain arms of several dozen base pairs in length, it is unlikely that the specific length is important since Chaires and Spink demonstrated that a 20 and 200 base pair duplex were stabilized to similar extents in the presence of macrosolute (32,74).
References


