DNA G-Quadruplex Formation in the *Bdellovibrio bacteriovorus* Genome: An *in vitro* Study Exploring Temperature, Time, and Crowding

Hikari Murayama
Advisor: Megan E. Nuñéz
Departments of Chemistry and Physics

Submitted in Partial Fulfillment of the Prerequisite for Honors in Chemical Physics
April 2016
© 2016 Hikari Murayama
This paper was prepared
under the direction of
Professor Megan E. Núñez
for 2 credits
Acknowledgements

I have the utmost gratitude for my thesis advisor, Professor Megan E. Nuñez. What drew me in initially into her lab is her utmost enthusiasm, love and excitement for science. Not only is she excited to teach you new things, she is excited to learn new things with you every step of the way. She knew almost nothing about me or what my individual major consisted of, but her flexibility, genuine willingness to help, and her supportiveness has really helped me grow the past two years I have spent researching in her lab. Furthermore, I thank my thesis committee members Professor Mala Radhakrishnan and Professor Jerome Fung for guiding me through the process and always giving me suggestions; I look up to you both and I admire your thoughtfulness, kindness, and your patience. I also have to thank Professor Paul Reisberg and Professor Kača Bradonjić for being honest mentors in addition to good discussions on everything but chemistry and physics.

My project could also not have been completed without the collaborative efforts from my fellow lab mate and thesis mate, Sally Shepardson-Fungairíño. Without her help I never would have been able to kick start this project. And thank you to our other lab mates for tolerating us and being nice to us when you knew the CD wasn't: Leah Furman, Maggie Kluerza, Lucy Ortega, Sally Ruderman, and Divya Satischandra.

Both Sally S.F. and I have many, many thanks for George Dai and Elaine Igo have been spectacular in helping us acquire such beautiful spectra. Thank you for the countless hours on the phone with the Olis company, us emailing and calling you to figure out how to solve the endless issues that came with us using the machine. We are so
grateful that we have such patient and knowledgeable instrument specialists at Wellesley College. We must also thank John DeBolt for the kind help as well.

And where would I be without my lovely friends that have supported me throughout the whole process. From the endless nights in the Science Center (where Professor Nuñez jokes that we should just leave a “reserved for Chemistry Majors” sign), to the fun nights off hanging out in the dorms, I cannot fathom my college experience without you. This is to you Jane Zhu, Darlene Dang, Amy Yuan, and Hannah Sim. (Cheers to my other Chemistry and Physics majors too!) And of course, thank you to my non-science center hibernating friends: Caroline Chang, Wanyi Li, Rebecca Chen, Crysti Wang and my Japan Club family.

Last but not least, I am far too in debt to my family. My siblings, for always engaging in my quibbles, and competing against one another in every subject we’ve shared learning about; I hope that you will continue to support me and let me continue to support you throughout all of our endeavors. My parents, you are the most forgiving, understanding and encouraging parents I know. Even though we don’t talk all the time, thanks for always figuring out ways to express your kindness, not usually in words, but in gestures.

I could not have completed this project without all of the support.
Table of Contents

List of Tables ............................................................................................................. vi
List of Figures ............................................................................................................. vi

1. Abstract .................................................................................................................. 1

2. Introduction ............................................................................................................ 3
   2.1 DNA G-Quadruplex .......................................................................................... 3
   2.2 Ultraviolet to Visible (UV-Vis) Spectroscopy Theory ...................................... 8
   2.3 Circular Dichroism (CD): Theory .................................................................. 9
   2.4 Circular Dichroism Instrumentation ................................................................. 16
   2.5 Molecular Crowding ....................................................................................... 19
   2.6 DNA Sequence Selection ................................................................................. 24
   2.7 Purpose of Study ............................................................................................. 25

3. Methods ................................................................................................................ 26
   3.1 Sample Preparation ......................................................................................... 26
   3.2 Circular Dichroism Measurements ................................................................ 28

4. Results .................................................................................................................. 29
   4.1 Time vs. Temperature ...................................................................................... 29
   4.2 Crowding Experiments: PEG 600 .................................................................. 35
   4.3 Crowding Experiments: PEG 8000 ................................................................. 43
   4.4 Crowding Experiments: Ethylene Glycol ....................................................... 50

5. Discussion ............................................................................................................ 56
   5.1 Time vs. Temperature ...................................................................................... 56
   5.2 Crowding Experiments .................................................................................... 67

6. Conclusion ............................................................................................................ 76

7. References ............................................................................................................ 79
List of Tables

Table 1. Summary of characteristic CD spectra peaks and their signs for different quadruplex conformations ................................................................. 13
Table 2. Energetics of human telomere quadruplex unfolding ............................................... 61
Table 3. G-quadruplex stabilities under molecular crowding conditions .................................. 74
Table 4. Hydration of G-quadruplexes ................................................................................. 74

List of Figures

Figure 1. G-quartet 4 guanines in one or several DNA strands ........................................... 4
Figure 2. Unimolecular, bimolecular and tetramolecular G-quadruplexes ............................... 5
Figure 3. Quadruplex and its inhibitory properties of biological processes ......................... 7
Figure 4. Probable transition dipoles for Guanine ............................................................... 9
Figure 5. Coronene, an example of an achiral molecule ....................................................... 10
Figure 6. The syn and anti glycosidic bond angles of guanosine ........................................... 11
Figure 7. Two forms for G-quartet stacking ........................................................................ 12
Figure 8. Typical CD spectra of standard structures .............................................................. 15
Figure 9. Optical path in Olis RSM 1000 Rapid-Scanning Spectrometer with parallel PMTs for CD .................................................................................. 18
Figure 10. Schematic of the Olis RSM 1000 Rapid-Scanning Spectrometer ......................... 18
Figure 11. Molecular environment inside the cell ................................................................. 19
Figure 12. Four different effects caused by different types of crowding agents ................. 20
Figure 13. CD Spectrum of sample in 5mM KCl, 0.01 M Tris buffer. ................................. 32
Figure 14. CD Spectra on Day 0, Day 4, Day 7, and Day 11 ............................................. 33
Figure 15. Change in CD spectrum over time at 20°C and at 10°C for sample at 5 µM in 0.01 M Tris buffer ................................................................. 34
Figure 16. CD spectrum of sample in 5mM KCl, 0.01 M Tris buffer and 10% (w/v) PEG 600 ...................................................................................................... 38
Figure 17. CD spectrum of sample in 5mM KCl, 0.01 M Tris buffer and 20% (w/v) PEG 600 ...................................................................................................... 39
Figure 18. CD spectrum of sample in 5mM KCl, 0.01 M Tris buffer in 30% (w/v) PEG 600 ...................................................................................................... 40
Figure 19. CD spectrum of sample in 5mM KCl, 0.01 M Tris buffer and 40% (w/v) PEG 600 ...................................................................................................... 41
Figure 20. CD spectrum of sample in 5mM KCl, 0.01 M Tris buffer and 40% (w/v) PEG 600 ...................................................................................................... 42
Figure 21. CD spectrum of sample in 5mM KCl, 0.01 M Tris buffer and 10% (w/v) PEG 8000 ...................................................................................................... 46
Figure 22. CD spectrum of sample in 5mM KCl, 0.01 M Tris buffer and 20% (w/v) PEG 8000 ...................................................................................................... 47
Figure 23. CD spectrum of sample in 5mM KCl, 0.01 M Tris buffer and 30% (w/v) PEG 8000 ...................................................................................................... 48
Figure 24. Comparing CD spectra of 40% (w/v) PEG sample with complement ........ 49
Figure 25. CD spectrum of sample in 5mM KCl, 0.01 M Tris buffer and 10% (w/v) EG .............................................................. 52
Figure 26. CD spectrum of sample in 5mM KCl, 0.01 M Tris buffer and 20% (w/v) EG ........................................................................ 53
Figure 27. CD spectrum of sample in 5mM KCl, 0.01 M Tris buffer and 30% (w/v) EG. ................................................................. 54
Figure 28. CD spectrum of sample in 5mM KCl, 0.01 M Tris buffer and 40% (w/v) EG. ................................................................. 55
Figure 29. CD spectra of G3 in 10 mM Cs-HEPES buffer, 100 mM K+ at pH 7.5 at 20°C by Olsen et al.................................................. 57
Figure 30. CD spectra of G3 sequence in 10mM Cs-HEPES buffer, 100 mM K+ at pH 7.5 from 20°C to 105°C by Olsen et al........................ 58
Figure 31. Isothermal folding kinetics for Rehm et al study. ........................................ 59
Figure 32. CD spectrum of 0.01 M Tris buffer blank.............................................. 64
Figure 33. CD spectrum of 40% PEG 600 Blank...................................................... 65
Figure 34. CD melting curves for 5µM DNA sample in 0.01 Tris Buffer at 265 nm and at 290 nm ........................................................................ 66
Figure 35. CD spectra of G3(T2AG3)3 in 150 mM K+ solution and 150mM Na+ solution with various concentrations of PEG. ......................................................... 69
Figure 36. CD spectra of various crowding agents at 42.5% (v/v) concentrations, with another plot showing showing the titration curves of each crowding agent that was used at 290 nm ..................................................................... 70
Figure 37. CD spectra of 50 µM of intra-Tet and intra-Hum at 4°C. ....................... 72
1. Abstract

DNA G-quadruplexes form when an abundance of guanines in a sequence interact via Hoogsteen bonding to create a stable secondary structure. This particular conformational type is of interest because of its abundance in telomere sequences, promoter and immunoglobin switch regions, potentially affecting transcriptional and translational processes. In this study we used a sequence from the *Bdellovibrio bacteriovorus* genome, 5'-GGGTTTGGTGTCAGGGGCAGGG-3’, a sequence that encodes a gene for a cell division protein *fts*.

Using an HPLC purification protocol that is designed for G-quadruplex forming DNA sequences, the DNA oligonucleotide was purified of any side products. After dialysis the DNA was diluted in 0.01 M Tris buffer (pH 7.3). These were analyzed by circular dichroism spectroscopy (CD). The ellipticity was measured from 230 to 340 nm, spanning temperatures from 10°C to 90°C.

The folding of our oligonucleotides showed marked dependence on both time and temperature. Subsequent scans of the same sample exhibited different antiparallel conformation mixtures; furthermore, constant temperature scans over a prolonged period of timed also showed signal changes.

The crowding studies used the polymeric agents polyethylene glycol 600 (PEG 600) and PEG 8000, as well as the small monomeric ethylene glycol (EG). We saw differences in spectra with length and concentration of crowding agent. Increasing PEG 600 concentration caused a conformational shift for the whole spectrum from a mixture of group II and group III antiparallel quadruplexes to a group I parallel quadruplex system when the crowding concentration was increased to 40% (w/v). PEG 8000 also
showed a similar pattern, though the 40% (w/v) concentration showed spectra similar to that of the complement sequence. EG did not exhibit significant change as compared to the other crowding agents.

These results indicate the critical role of time and temperature in developing a robust quadruplex protocol. In the future, differential scanning calorimetry (DSC) can be used to find kinetic constants to complement this work and isolate the effects of time and temperature. By continuing this study, we will be able to apply our knowledge to in vivo quadruplex formations in Bdellovibrio bacteriovorus.
2. Introduction

2.1 DNA G-Quadruplex

While usually Deoxyribonucleic acid (DNA) folds into a double helix, other forms such as single-, triple- and tetra-stranded helices are also often present. When there is an abundance of guanines, tetrads form with Hoogsteen hydrogen bonds replacing the usual Watson-Crick base pairing (Figure 1).\textsuperscript{9,10} Since this type of structure is found in telomeric sequences, promoter and immunoglobin switch regions, quadruplexes have an effect on transcriptional and replication processes.

These secondary structures have been located \textit{in vivo}, tending to cluster in particular areas of the genome. They have been evolutionarily conserved within the human with homologs in yeast genomes and are thus thought to have a positive function in the cell. Quadruplexes are quite common in a variety of organisms at the ends of chromosomes, called telomeres, due to the high guanine and cytosine content and the single strand overhang characteristic of telomeres.\textsuperscript{6}

The 5’ end of eukaryotic chromosomes consists of a repetitive thymine guanine sequence, while the 3’ complement consists of an adenine cytosine sequence. The number of bases per type is usually from 1 to 4, (i.e. (dT\textsubscript{1-4}dG\textsubscript{1-4})\textsubscript{n}) and these sequences can repeat from 20 to 100 times in a single celled eukaryote and 1500 times for mammals. These sequences are added to the end of the chromosome by telomerase, which is a ribonucleoprotein that synthesizes telomeric DNA by using its RNA component as a template.\textsuperscript{6}
A G-tetrad characteristically is formed by 4 guanines from different strands (intermolecular) or from the same strand (intramolecular) interacting by Hoogsteen bonding to form a guanine tetrad. Since the N₁ and N₂ are hydrogen bond donors, while O₆ and N₇ are the hydrogen bond acceptors, the four guanines can associate in a square planar form as illustrated in Figure 1. Since the carbonyl O₆ is electronegative, a cation, such as K⁺ and Na⁺, fits between the bases to relieve those charge repulsions. When stacked upon one another, multiple g-tetrads form a quadruplex with the interspersing non-guanine nucleotides forming single-stranded loops of various lengths. Because of the stability associated with base stacking interactions between adjacent tetrads, common sequence motifs have at least 3 guanines in the tract. Loops are typically one to seven nucleotides in length, with smaller loops resulting in more stable structures. These strands can either be all parallel to one another, antiparallel to one another, or a combination hybrid of the two (Figure 2).
Figure 2. Unimolecular, bimolecular and tetramolecular G-quadruplexes. The top two rows show the probable structures of propeller, chair, basket and 2 types of hybrid conformations for a single strand DNA. The next row shows the three possible structures for a bimolecular structure, which are lateral, diagonal and double-chain reversal loops. The tetramolecular formation is shown last.3
Because the 5' ends are not readily replicated by DNA polymerase, without telomerase the chromosomes would be shortened in each cell generation, losing about 50 to 200 nucleotides per cell division. In fact, most mature cells downregulate the gene for telomerase while 80 to 85% of tumor cells express it, enabling cancer cells to divide indefinitely.\textsuperscript{11} G-quadruplexes are of interest here since they are able to inhibit telomerase activity. Small molecule ligands, as seen \textit{in vitro}, can bind to target regions to help stabilize quadruplex structures to inhibit excessive cell generation.\textsuperscript{6}

Quadruplexes also have an effect on DNA replication since they can form in double-stranded DNA, especially in the promoter and immunoglobin switch regions.\textsuperscript{12} Replication involves the double helix separation by replicative helicase, in which each strand has a specific characteristic: the lagging strand is replicated in segments, while the leading strand is replicated continuously. Because of the block-by-block replication of the lagging strand, it is characteristically single stranded, allowing opportunities for intramolecular quadruplex formation. This formation may slow down progress of DNA polymerase, and may also allow the regulation of transcription. Without the unraveling of these structures, likely by helicases, replication cannot be completed (Figure 3a).\textsuperscript{6}
Since promoter regions of genes can have a locally high concentration of guanines, it is believed that quadruplexes may also have an effect on transcription. For example, in 50% of the human genome one or more guanine tracts are within 1000 nucleotides upstream of transcriptional start sites (TSS). Supercoiling in these regions can induce both positive and negative effects. Thus, in the event of abundant guanine content, quadruplexes may form, accelerating or decelerating transcription. In the case that the guanine rich sequence is encoded on the template strand, transcription would not be able to proceed (Figure 3b). On the other hand, if it is located on the non-template strand, it can allow the easy transcription of the now single stranded template strand.

**Figure 3. Quadruplex and its inhibitory properties of biological processes**

(a) Unresolved quadruplexes can inhibit DNA replication by slowing or stalling the replication fork sterically by blocking the replisome (blue). (b) Quadruplexes can also inhibit transcription by acting as a physical block for RNA polymerase (purple).
Furthermore, proteins bound to the quadruplex structure can also enhance or inhibit transcription depending on whether the protein is an enhancer or a suppressor.\(^6\)

### 2.2 Ultraviolet to Visible (UV-Vis) Spectroscopy Theory

When measuring absorption, we determine the amount of a light with a certain frequency \((\nu)\) that is absorbed by a sample. When this occurs, the absorbed photon increases the energy of the sample in question by the following relationship:

\[
\Delta E = h\nu = \frac{h\omega}{2\pi}
\]

in which \(h\) is Planck’s constant while \(\omega\) is the angular frequency. A polarization of transition occurs, where the electron density is pushed by the electric and magnetic field to a higher energy, producing a transition moment. However, we do not get a sharp peak since the molecules are all of slightly different energies. Multiple explanations for this distribution include the Uncertainty Principle \((\Delta E\Delta t \geq \frac{\hbar}{4\pi})\), Doppler broadening, intermolecular interactions, rotational transitions, and vibrational transitions.\(^{14}\)

Nucleic acids are studied in the range 200 to 300 nm because of transitions associated with the purine and pyrimidine bases. As illustrated in Figure 4 for guanine the assigned \(\pi \rightarrow \pi^*\) transition dipoles have an assigned axis. Since the ribose-phosphate backbone has shown to have no significant transitions in this range, it is understood that we are measuring either the transition of the base coupling with the backbone, or in the case of quadruplexes, the transition of the base coupling with another nearby base. Differentiating between these two scenarios is possible since the transition for the former is around 200 nm while for stacked bases the absorption is detected close to 270 nm.\(^{14}\)
Circular Dichroism (CD): Theory

Circular Dichroism (CD) is a technique that uses the differences in absorption of right- and left-handed circularly polarized light to understand the properties of chiral molecules. Two plane-polarized waves, oriented perpendicular to one another, can be superpositioned out of phase relative to each other (\(+\pi\) or \(\pi\) radians) to produce circularly polarized light. In this case the overall vector on a certain \(xy\) plane would rotate in a circle with a period of \(2\pi\) with a constant vector length. The superposition of the two circularly polarized lights is not as trivial as that of two plane-polarized lights. Take the case of a superposition of an in phase right- handed and left-handed circularly polarized light, and assume that they are of the same wavelength and amplitude. In this case, the superposition would result in plane-polarized wave.

When light interacts with matter, several characteristics are susceptible to change: intensity (amplitude), velocity (due to the refraction index of the medium), polarization, and wavelength. For CD, the important concept is absorption, which entails that the

---

**Figure 4. Probable transition dipoles for Guanine.** When coupled with the backbone, the guanine shows two distinct transitions. This base acts as the chromophore to detect the CD signal. Adapted from Rodger et al. \(^1\)
matter absorbs part of the light because of characteristics related to chirality and dipole moments. The importance lies in the difference in absorption of left circularly polarized light versus that of the right circularly polarized light. When the matter being transversed absorbs one direction of light more than the other, the result is no longer plane-polarized but elliptically polarized.\textsuperscript{16} The quantity of the difference between the absorbance is defined as the ellipticity ($\theta$).\textsuperscript{15}

In order to understand CD, optical spectroscopy needs to be understood first on a fundamental level. With the electric and magnetic fields produced by the electromagnetic radiation, there is an excitation of the electrons to a new stationary state; each component serves a different role in this change. The electric dipole transition moment (edtm) or $\vec{\mu}$, which is a linear rearrangement of the electrons, is caused by the electric field and points towards the direction that the charge has been repositioned to. On the other hand, the magnetic dipole transition moment (mdtm), or $\vec{m}$, is a result of the magnetic field, as the name implies, and causes a circular rearrangement as opposed to the linear one of the edtm. So, for an achiral molecule, such as coronene (Figure 5), the electron distribution is planar and depending on whether it is linear or circular, can have $\vec{\mu} = 0$, $\vec{m} = 0$, or both. However, a chiral molecule will have a helical electron rearrangement, in which the net $\vec{\mu}$ and the net $\vec{m}$ are parallel. This phenomenon is what is detected on the CD in terms of a positive and

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{coronene.png}
\caption{Coronene, an example of an achiral molecule. Since the electron distribution is planar and the molecule itself is circular, is the case that both $\vec{\mu}$ and $\vec{m}$ are 0.}
\end{figure}
negative peak. For instance, if the alpha helix is right-handed, the excitation is more easily induced by a left circularly polarized light, creating a positive CD signal. Conversely, if the helix is left-handed, the excitation is induced by a right circularly polarized light, creating a negative signal. This is shown concisely within the following Rosenfeld equation for CD rotational strength for randomly oriented chiral molecules:

\[ R = \text{Im}[\mu \cdot m] \]

in which Im denotes the imaginary portion, \( \mu \) is the edtm of the transition from final to initial, and \( m \) is the mtom for the reverse transition. Therefore to have a CD signal, the sample must have an electronic and magnetic transition.

In the case of a G-quadruplex, the differences in CD spectra are highly reliant on the folding topologies and whether the structure incorporates syn or anti glycosidic bond angles (GBA) within the guanine repetition (Figure 6). Recall the possibility of a parallel, antiparallel or hybrid conformation. With these different orientations, the structure will have different relative strand directions, groove widths and loop orientations. Furthermore, the stacked G-tetrads can have either a heteropolar (head-to-head or tail-to-tail) stacking or homopolar stacking (head-to-tail).
When the chromophores are coupled and have a distinct allowed transition, the CD spectrum intensity is able to be informative about the quadruplex structure.\textsuperscript{17} A chromophore can be defined as any light-absorbing molecule or group, which in our case comes from the purine and pyrimidine bases in the DNA.\textsuperscript{18} A chromophore should be a distinct part of the system that does not have an overlapping wavefunction with the rest of the system, and thus for CD measurement requires that the portion has an absorption that is distinct and independent from the rest of the system under study.\textsuperscript{14}

If two chromophores are coupled and are chirally oriented in opposite directions relative to one another, the spectrum will exhibit two bands of opposite sign, and their overlap $\lambda_{\text{max}}$ will have an intensity of 0. In the quadruplex case, we look at $\lambda > 210$ nm where the chromophores are the guanines, and the rest of the DNA strand is negligible. These have two well-defined absorption bands between 240 and 290 nm and have

\textbf{Figure 7. Two forms for G-quartet stacking} The left shows homopolarity stacking in which there is head to tail association between the two stacking guanines. On the other hand, there is heteropolarity stacking, which is shown on the right. The two vectors show the transition dipole at ca. 250 nm.\textsuperscript{17}
characteristic $\pi-\pi^*$ bands at 248 and 279 nm for the transition of the base coupled to the backbone (Figure 6). The rotation of the stacked bases causes excitation of the chiral coupling with neighboring guanine transition dipole moments. There are two ways the G-quartets can stack on top of each other: homopolar and heteropolar. Homopolar involves a head to tail stacking, while heteropolar involves a head to head stacking. When there are electronic allowed transitions, as in this case of $\pi-\pi^*$, it is largely due to the electric dipole transition while the magnetic dipole transition moment contributes only a negligible amount. The transition dipole moments for the stacked bases differ depending on whether it is H-to-T or H-to-H (Figure 7).\textsuperscript{17}

Each guanine in the stem in the quadruplex can adopt either an \textit{anti} or a \textit{syn} GBA, and Randazzo et al. has classified these systems into three different groups. Group I consists of all guanines with the same GBA with all stands parallel. The other two groups are characteristically anti-parallel. Group II involves antiparallel strands and has guanines of different types (\textit{anti-syn} and \textit{syn-anti}) as well as the same (\textit{anti-anti} and \textit{syn-syn}). Group III, although antiparallel, has consecutively stacked guanosines with a specific type of GBA.\textsuperscript{17}

Karsisiotis et al. characterized the topology of G-quadruplexes by CD; because of their decreases in etdm, or hypochromisity, a substantial difference can be detected using

<table>
<thead>
<tr>
<th>Group</th>
<th>Peak A (240 nm)</th>
<th>Peak B (260 nm)</th>
<th>Peak C (290 nm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group I: Parallel</td>
<td>Negative</td>
<td>Positive</td>
<td>0</td>
</tr>
<tr>
<td>Group II: Anti-Parallel</td>
<td>Negative</td>
<td>Positive</td>
<td>Positive</td>
</tr>
<tr>
<td>Group III: Anti-Parallel</td>
<td>Positive</td>
<td>Negative</td>
<td>Positive</td>
</tr>
</tbody>
</table>

Table 1. Summary of characteristic CD spectra peaks and their signs for different quadruplex conformations.
CD. Figure 8a shows that a spectrum of a typical Group I with three or four loops has a negative band at 245 nm with a positive band at 264 nm. Figure 8b shows a positive peak at 290 nm, which is characteristic of an antiparallel conformation, a negative peak at 240 nm and a positive band at 260 nm. The reverse is true for Group III, in that although it shows a positive peak at 290 nm, it has a positive peak at 240 nm while it has a negative band at 260 nm (Figure 8c). A summary of all of these characteristics can be found in Error! Reference source not found.. These observations translate to concrete conclusions about the GBA stacking: a band at 290 nm indicates the presence of stacking of different GBA guanosines; a positive band at 240 nm with a negative band at 260 nm indicates different or mixed GBA stacking in the quadruplex stem; and a negative band at 240 nm with a positive band at 260 nm is indicative of uniform GBA stacking.17

The measurement itself is differential absorbance in units of ellipticity (Θ), so to convert to absorbance (A) we can use the following equation:

\[ \Theta = \Delta A = A_L - A_R = \frac{4\pi \Theta \text{ (degrees)}}{180 \ln 10} = \frac{\Theta \text{ (millidegrees)}}{32,982} \] (3)

In addition, the Beer-Lambert Law for the CD can be written as:

\[ \Delta A = (\Delta \varepsilon)C\ell \] (4)

where \( \ell \) is the path length, \( \Delta \varepsilon \) is the difference in extinction coefficients, and \( C \) is the concentration of the sample.14
Figure 8. Typical CD spectra of standard structures (a) shows Group I members, which are parallel quadruplexes, with a negative peak at 245 nm, and a positive peak at 265 nm. Although Group II and Group III are both antiparallel, they have characteristically different spectra. (b) shows Group II with a positive peak at 290 nm, a negative peak at 240 nm and a positive peak at 260 nm. On the other hand, Group III in (c) has a negative peak at 260 nm, a positive peak at 240 nm and at 290 nm.
2.4 Circular Dichroism Instrumentation

The core schematic of the Circular Dichroism Instrument is illustrated in Figure 10. For the study presented in this thesis, an Olis RSM 1000 Rapid-Scanning Spectrometer was used, which consists of several different components: the cooling box, the lamp power supply and the Xe lamp, the monochromator, the PEM, and two PMTs. The signal detected via this instrumentation is directly transferred to the computer.\(^{20}\)

The UV-light is produced by a 150-watt xenon arc lamp powered with 7V. Since this lamp produces ozone when in use, a constant flow of nitrogen gas is required to remove oxygen from the apparatus. In addition, to make sure the lamp does not overheat, a flow of water is used as a coolant.\(^{21}\)

The monochromator allows the selection of certain wavelength of light for measurement. The Olis RSM 1000 Rapid-Scanning Spectrophotometer has a DeSa monochromator, a model that utilizes two halves that are mirror images of one another: one diffracts the initial light beam while the second recombines the first half to produce a homogeneous output beam.\(^{20}\)

Part of the output of the monochromator goes in the photo elastic modulator (PEM). This device is able to transform a linearly polarized light to that of a left or a right circularly polarized light, so that the two beams are now \(\pi\) radians out of phase. The installed model in the CD used in this project transforms this light from linear to circular at a rate of 50 kHz.\(^{20}\)

The light from the PEM next passes through the sample. Both right and left circularly polarized light are passed through the chosen sample simultaneously, and the result is detected by two photo multiplier tubes (PMT). The cathodes and anodes within
the PMTs are able to multiply the light detected and are highly sensitive to wavelengths in the UV region. 

The signal is then received by the MCB and computer. The MCB contains a twin 14-bit analog to digital (A/D) converter board that is able to convert the analog signal into a digital one to finish the data collection scheme.
Figure 10. Schematic of the Olis RSM 1000 Rapid-Scanning Spectrometer The instrument consists of several different parts. The cooling box which cools the lamp, the lamp power supply and the Xe lamp, the monochromator that adjusts the light to various wavelengths, the PEM, which then shines through the sample. The light is then detected by the PMTs and the signal is converted to a digital one from an analog, and is then able to be detected by the computer.  

Figure 9. Optical path in Olis RSM 1000 Rapid-Scanning Spectrometer with parallel PMTs for CD
2.5 Molecular Crowding

The cell consists of macromolecules of proteins, nucleic acids, carbohydrates. By analysis using electron microscopy and fluorescence spectroscopy researchers were able to visualize shapes, distributions, stability and dynamics of various components of the cell. The studies conducted by Nakano et al. and Miyoshi et al. revealed that 30-40% of the total volume of the cell is occupied by biomolecules, with a total concentration of about 400 g/L.\textsuperscript{7,22}

In order to make the biochemical experiments relevant to \textit{in vivo} conditions, the crowded nature of the cell created by various organelles and macromolecules also has to be mimicked. To tackle this, many groups have applied these molecular crowding techniques by using various “inert” chemical agents, which fall into several different categorical types of crowding. Macromolecular crowding occurs when large molecules, such as proteins and nucleic acids, create the crowded condition. These result in repulsive

\textbf{Figure 11. Molecular environment inside the cell} Contrary to how the cell may be introduced in introductory biochemistry and biology courses, the inside of the cell is crowded with high concentrations of all of its components.\textsuperscript{7}
interactions and obstacles for the movement of components of the cell. On the other hand, small-molecule crowding occurs when a large number of small molecules crowd the cell, which most often effects water activity and the hydration of other molecules. In this study we have defined water activity as inversely related to osmotic pressure or stress. So if there is an increase in water activity, there is a decrease in osmotic pressure and vice versa.

Several effects result from crowding. First there is the excluded volume effect, which excludes volume and makes it inaccessible for other components within the system or can also result from spatial confinement of the system. Next, since larger molecules can act as obstacles, crowding can also decrease the diffusion rate or the efficiency of molecular collisions. For these circumstances, the competition between enthalpy and

**Figure 12. Four different effects caused by different types of crowding** agents (a) The excluded volume effect cause the formation of inaccessible areas that act as obstacles. (b) With obstacles, a decrease of the diffusion rate can also be observed. (c) Some crowding agents are able to change the properties of the solution itself. One example is the dielectric constant. (d) The properties of the crowding agent can also cause changes between the biomolecules, influencing their interactions. This leads to changes with rate and equilibrium constants.
entropy must be taken into account. Third, there can be a change in solution properties such as water activity, where the water can form ordered layers and become inactive. There is also a decrease in the dielectric constant when comparing pure water with that of a molecule rich solution, which causes more electrostatic interactions. Finally, the equilibrium constants and rates for intracellular reactions are dependent upon the environment in which they take place.\(^7\)

Ideal synthetic cosolutes for a crowding experiment need to meet the following characteristics: they are very soluble in water, they do not cause precipitation of nucleic acids, and they do not form stronger bonds to nucleic acids and metal ions than water. To adhere to these three criteria, neutral molecules such as polyethylene glycol (PEG) and polysaccharides are usually used.\(^7\)

Nonetheless, instances of precipitation as a result of using PEG and other interactions with nucleic acid that can occur. In 1975, Lis and Schleif fractionated DNA by size using precipitation with PEG.\(^23\) Paithankar and Prasad obtained good precipitation and recovery of precipitated DNA using a PEG 8000 concentration of 13\% with 10 mM of MgCl\(_2\), at room temperature.\(^24\) Therefore, if using PEG as a crowding agent, the researcher must be careful to check for precipitation within their sample tube, although in this case some of the precipitation may be due to the large concentration of Mg\(^{2+}\).

Multiple groups have used PEG as a crowding agent for their quadruplex studies. Zheng et al. used DMS footprinting and gel electrophoresis to look at G-quadruplex formations in long double stranded DNA from the human genome and compared this in dilute and PEG-induced molecular crowding conditions. The group compared
quadruplex and duplex formation by heat denaturation and renaturation in both conditions. Fluorescence was also used to compare the stability of quadruplex vs duplex with the presence of PEG. They saw that with an increase in percentage of PEG 200, hairpin melting temperature decreased, while transition temperature for quadruplexes increased.\textsuperscript{25}

Xue et al. took this one step further by examining how telomeric sequences in DNA form parallel-stranded intramolecular G-quadruplex structures in K\textsuperscript{+} solutions under molecular crowding conditions.\textsuperscript{1} These conditions are important to understand since the intracellular milieu is characterized by high K\textsuperscript{+} concentrations, around 150 mM. They demonstrated that both the crystalline and solution state have the same conformation using three different methods: CD, gel electrophoresis, and fluorescence. Recall that CD has a couple key characteristics to note. For an antiparallel structure there is a negative peak at 265nm (peak B) and a positive peak at 295 nm (peak C). For a parallel structure, there is a negative peak at 240 nm (peak A) and a positive peak B. Their spectra showed an increase of a peak B and a decrease in peak C, demonstrating a change in conformation from an antiparallel one from a parallel one with a larger crowding agent concentration. Miyoshi et al. found that PEG crowding induces a quadruplex transition from antiparallel to parallel, while polycations do not cause conformational shifts. They claim that the transition is entropy driven and that electrostatic interactions and volume exclusion are the key drivers.\textsuperscript{22}

Other groups used computational methods to look at molecular crowding and its effect on G-quadruplex formations. Doghaci et al. used MD simulations to determine the effects of molecular crowding by ethanol with the presence and absence of K\textsuperscript{+} cations.
They determined that the G-quadruplex collapses in water in the absence of potassium, with ethanol stabilizing the structure by volume exclusion and decreasing water activity. In addition, ethanol can increase the stability of the Hoogsteen bonding in the G-tetrad.26

However, it is also important to note that each group looked at slightly different DNA sequences. Karimata et al. argued in their study Tetrahymena (Tet) and human (Hum) telomere sequences that a single base difference can lead to entirely different structures under molecular crowding solutions.4 For intra-Tet they observed a positive peak C and a negative peak B without co-solutes in the CD spectra. With high molecular weight crowding co-solutes, the peak B signal increased, indicating a transition from antiparallel to a parallel conformation. This led to the conclusion that a transition from a monomeric antiparallel G-quadruplex to a multi-stranded G-wire can occur. However, intra-Hum did not show such a transition. The melting curves in the paper are also exemplary of this difference, where the intra-Tet Tm increased remarkably with an addition of 40% (w/v) co-solute, while intra-Hum increased only slightly, with higher weight co-solutes causing the largest deviation from the original signal.

Although multiple studies have concluded that PEG causes a transition from anti-parallel hybrid to an all-parallel quadruplex conformation, it is unclear why crowding should cause this transition.13 Buscaglia et al. claim that other agents such as BSA, Ficoll 70, and Ficoll 400 do not cause this structural change. For a true crowding simulation, the perturbation of the equilibrium should not be reliant on the agent used. Therefore, they conclude that PEG is not altering the conformation by crowding but by some other unknown mechanism that they propose are binding interactions. After many in vitro experiments looking at water activity and excluded volume effects, they also realized that
the propeller conformation must be larger than hybrid forms and also confirmed their assertion by molecular dynamics.\textsuperscript{8}

Buscaglia et al. similarly argued that PEG is not a good crowding agent; rather, it binds to the G-rich sequences, telomeric sequences in this case, and therefore is not a good agent to mimic intracellular crowding. It is generally agreed that water activity is changed by adding small co-solutes, while adding large polymers act as obstacles, but in both cases it is essential that these added solutes are inert and do not interact with either reactants or products of the reaction under study. In conclusion, the group claims that decreased water activity stabilizes quadruplexes and increases their melting temperature. To put it another way, formed quadruplexes are less hydrated than the single unfolded strands, thus pertains to how water is released when the tetrads are formed.\textsuperscript{3}

Although in this study we are considering only supposedly inert crowding agents, this does not indicate the complete picture within the cell. Considering \emph{in vivo} circumstances, one must realize there are also “sticky” crowding agents such as protein that the DNA sequences may interact with.

\textbf{2.6 DNA Sequence Selection}

The DNA sequence of interest for this project was selected from a genome wide analysis of \textit{Bdellovibrio bacteriovorus}, performed by Dr. Shantanu Chowdhury of the Institute of Genomics and Integrative Biology (Delhi, India). For intramolecular quadruplex structures, they set the criteria that there must be 4 runs of guanines, interspersed with the other three types of bases that would form the loops. Their Java algorithm also made several other assumptions:
1. The stem or run size is constant in a single quadruplex and is between 3 to 4 nucleotide bases.
2. The stems only consist of guanines.
3. The loop is limited to 5 nucleotide bases in length.
4. A quadruplex consists of 4 stems and 3 loops.

The analysis was thus able to detect the possibility of quadruplex formations both strands by searching through the entire genome. Among the sequences tested, we chose strands that tested positive for quadruplex formation and that were biologically interesting (i.e. within a gene or promoter). Among those, this paper focuses on the strand 5’-GGGTTTGGGTCAGGGGCAGGG-3’, a sequence that encodes the gene for a cell division protein ftsw.

2.7 Purpose of Study

The purpose of the study presented in this paper is to explore the effects of the variables time, temperature and crowding on the DNA sequence, 5’-GGGTTTGGGTCAGGGGCAGGG-3’, found in the *Bdellovibrio bacteriovorus* genome.

There have been many studies indicating the importance of all of these variables on the stability of structure, what conformation the sequence takes, as well as what mixture of conformations may be present. The importance to control all of these variables, as well as their relative importance were explored. In addition, we also found the importance of sample storage and handling. We hoped that by seeing multiple variables being tested, we would be able to develop a rigorous protocol which would allow us and other groups to have data that is reproducible as well as be able to explain the folding and unfolding phenomenon by using thermodynamic as well as kinetic understanding of the system. We sought to make this initial step by analyzing our samples using CD spectroscopy.
3. Methods

3.1 Sample Preparation

The DNA strand of interest, 5'-GGGTTCGGGCAGGGCGCCAGGG-3', was chosen from a study by the Chowdhury group of the Institute of Genomics and Integrative Biology, which indicated that the strand potentially has quadruplex forming regions. This sample was purchased from Integrated DNA Technologies (IDT, Coralville, IA). The DNA strand was re-suspended in 400µL of 0.01 M Tris buffer (pH= 7.3). After centrifuging the sample briefly on the Fisher-Scientific Mini Centrifuge to insure full recovery of the sample, the solution was transferred into 0.45 µM centrifugal filter units, with each sample split into two 250 µL units. It was filtered until all solution had passed through the filter. The sample was now ready to be purified by high-performance liquid chromatography (HPLC).

The Agilent Technologies 1260 Infinity HPLC was used in conjunction with the DNAPac© PA-100 (9x250mm), a 10 µL injection loop, and the associated software program. All prepared buffers were filtered to inhibit bacterial growth. Before sample purification, the instrument was cleaned thoroughly first with water and subsequently with acetonitrile and elution buffer. The purification protocol with ion exchange was adapted from the DNAPac© manual for this project. To purify the DNA samples, line A was put into a 25 mM NaOH solution and line B into a 1M NaCl in 25mM NaOH solution. The run was set up as shown in the table at right.

<table>
<thead>
<tr>
<th>Time (min)</th>
<th>%A</th>
<th>%B</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.00</td>
<td>90.0</td>
<td>10.0</td>
</tr>
<tr>
<td>2.00</td>
<td>50.0</td>
<td>50.0</td>
</tr>
<tr>
<td>20.00</td>
<td>10.0</td>
<td>90.0</td>
</tr>
<tr>
<td>20.50</td>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td>24.50</td>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td>25.00</td>
<td>90.0</td>
<td>10.0</td>
</tr>
<tr>
<td>40.00</td>
<td>90.0</td>
<td>10.0</td>
</tr>
</tbody>
</table>
25 µL of the dilute sample was loaded into the syringe as a test to determine the elution time. The indication consists of a peak at 260nm and a sharp peak on the mAU vs time plot by the DAD1. After determining the elution time, the remaining sample was purified in two runs, collected in a conical tube, and stored after addition of 5% (v/v) 0.1 M TE buffer (Tris and Ethylenediaminetetraacetic acid). The HPLC was cleaned in the opposite steps of that of the pre-clean.

The next step was to dialyze all samples against 5 L of 10 µM ammonium acetate using the GE Healthcare Mini Dialysis kit with a 1 kDa cut-off, 25-µL size or Spectrum Labscope Dialysis Tubing with a 1 kDa cut off depending on sample size. The dialysis buffer was changed every 24 hours over the course of 3 days to insure removal of the NaOH and NaCl from the sample. The entirety of the dialysis was completed in a cold room at 4.0°C. These samples were then frozen, either with a conventional freezer or dry ice, and then lyophilized overnight on the BenchTop Pro with Omnitronics. The result was then re-suspended in 0.01M Tris buffer (pH 7.3).

The samples were then concentrated by centrifugation it in an Amicon filter with 1 kDa MW cut off, inserting the filter so that it is perpendicular to the hinge. Samples were centrifuged 5000 xg for 30 minutes. Samples were dialyzed against 0.01M Tris buffer pH 7.3 for 72 hours, changing the buffer twice.

The concentration of the purified DNA was evaluated using UV-VIS spectroscopy on the Agilent Technologies Cary Series UV-Vis Spectrophotometer and taking the absorbance at 260 nm. Using Beer’s Law (A = εcℓ) and the extinction coefficient provided by IDT (206,300 L/mol·cm), we calculated the concentration. Samples were diluted with 0.01 Tris buffer pH 7.3 to a 10 µM strand concentration.
For crowding studies, the crowding agents polyethylene glycol (PEG) 600, PEG 8000, and ethylene glycol were used. Since the most biologically relevant concentration is 40% (w/v), the agent was pipetted in at concentrations of 10%, 20%, 30% and 40% prior to the CD runs.

3.2 Circular Dichroism Measurements

Circular Dichroism (CD) spectra were measured on Olis RSM 1000 Rapid-Scanning Spectrophotometer with a Quantum Northwest temperature controller and a 7.00 V supplied to the Xenon lamp. 1500 µL of the diluted samples with a DNA concentration of 10 µM were pipetted into a 5 mm path length quartz cuvette (Starna). A temperature script was used to control the temperature per cycle, collecting data at 10, 20, 30, 45, 60, 75, and 90°C on the upwards scan and at the same intervals on decreasing the temperature, conducting a total of 13 scans. The samples were incubated for 5 minutes at each temperature before taking data. The ellipticity at each selected temperature point was measured in 55 wavelength increments between 230 to 320 nm, with an integration time of 55 seconds per point.

Results were retrieved, converted to .xls files, and plotted via MATLAB, subtracting out the base line and smoothed. If data were collected in the range 230 to 340 nm, we subtracted out the offset from 340 nm, where there should be no signal, to bring the baseline of the entire spectrum to zero. The given ellipticity measurements can be converted to concentration-corrected universal units of ellipticity for CD using the following equation for quantitative analysis:

\[
Corrected \Theta = \Delta A \left(\frac{1}{1000}\right) \left(\frac{1}{[DNA]}\right) \left(\frac{1}{Path \ Length \ (cm)}\right) \left(\frac{1}{\# \ of \ bases}\right) \ast 3300
\] (5)
4. Results

4.1 Time vs. Temperature

To understand the stability and thermodynamic properties of quadruplex formations, CD spectra were obtained between 230 nm to 340 nm at a range of temperatures from 10°C to 90°C. Thus our spectra show the melting and reannealing of quadruplexes for our DNA sequence, 5’-GGGTTTGGGTCAGGGGCAGGG-3’.

Figure 13 shows a spectrum with our DNA at a 5 µM concentration in a 5 mM KCl and 0.01 M Tris buffer solution. For temperatures between 10°C and 45°C, there is a distinct negative peak at 240 nm (which we will now refer to as peak A), a positive hump at 260 nm (peak B), and a tall peak at 290 nm (peak C). However, as the temperature rises higher to 60°C and up to 90°C, the shape of the curve starts to alter. At 60°C, the intensity of peak C decreases to the point where it is comparable to the height of peak B. At 75°C, the shape of the curve changes with no peak C, although it does keep the negative dip for peak A, and a positive peak B, albeit the intensity for that peak B has decreased. At 90°C, the intensity of the curve decreases, marking the minimum for all intensities compared to all spectra taken at the lower temperatures.

Notably, we found that with multiple CD runs of the same sample, the spectra unexpectedly changed from run to run. Figure 14 shows the CD spectra of the same 5 µM sample in 0.01 M Tris buffer taken over a period of 11 days: day 0, day 4, day 7 and day 11. The shape of the spectra on day 0 is different than on the other three days. Day 0 clearly has a higher intensity for peak B compared to peak C, with the relative heights of the two peaks staying the same as the temperature is ramped from 20°C to 90°C and back down, but with an overall decrease in intensity for the peaks. The spectrum also
loses its distinct humps as this spectrum is re-taken with increasing time and temperature. By comparison with the claims by Karsisiotis et al., because of the relative height difference in the two peaks and a negative peak A, we propose that the DNA sample may include a combination of group I parallel complexes and group II antiparallel complexes.

Day 4 through 11 all show similar shapes and patterns. At the initial temperature, all 3 spectra show a negative peak A, a positive peak B and a positive peak C. The temperature ramp up shows a high intensity peak C, compared to peak B, while the ramp down has a larger peak B and a smaller peak C. Instead of staying as a mixture of group I and group II complexes, the system is transitioning to a system with more group I and less of the group II complexes.

Overall, we see that the system is not at equilibrium over the timeframe of the experiment. There is a loss in signal when heating the sample as expected when the structure melts, but the signal is only regained a small amount or none at all when recooling from 90°C to 10°C, indicating that it does not have sufficient time to re-anneal. Instead, the bulk of the reformation of the quadruplex structure occurs when the DNA is stored in the refrigerator. The study led us to conclude that the CD scans were done on too short of a timescale and that the quadruplex takes longer than 24 hours to reach an equilibrium state.

Most noteworthy of this dataset is the fact that an initial annealing step is missing (in contrast to as studies such as Miyoshi et al.), but after the initial spectrum, the subsequent spectra are almost identical every time the spectrum is retaken. Therefore, as the sample is stored in the refrigerator at 4°C, the system is transitioning from a larger to smaller ratio of group I to group II quadruplexes.
Figure 15 shows the CD spectra obtained on the sample over a 20 hour period at constant temperature, measured after the sample was stored in a fridge at 4°C. Figure 15a shows the spectrum taken at 20°C, while b shows that of taken at 10°C. For both cases, although both the peaks at 240 nm and 290 nm stay at relatively the same intensity, the peak at 260 nm decreases from 5 to 3 as the sample slowly “warms up” from the 4°C conformation to the 20°C conformation.

These data clearly show the critical roles of both t and T as parameters in determining quadruplex topology. They emphasize that proper annealing and slow, patient measurement are critical to determine the equilibrium state of the quadruplexes.
Figure 13. **CD Spectrum of sample in 5mM KCl, 0.01 M Tris buffer.** Wavelength ranges from 230 to 320 nm, with the temperature being ramped up and down from 10°C to 90°C. At lower temperatures (10 to 60°C), the spectra show a negative peak A, a positive peak B, and a positive peak C indicating a possible mixture of antiparallel quadruplexes. At higher temperatures, the spectra have a negative peak A and a positive peak B, indicating the possibility of a group I parallel conformation.
Figure 14. CD Spectra on Day 0 (top left), Day 4 (top right), Day 7 (bottom left), and Day 11 (bottom right) for the sample at 5µM in 0.01 M Tris buffer. Sample was stored at 4°C in a conventional refrigerator. The first spectrum exhibits a unique trend compared to that of the others and has a higher intensity peak B than at peak C. This could be a mixture of group I and group II quadruplexes. Day 4 through 11 show a higher peak C than peak B, but then reanneals to a conformation similar to that of Day 0. These could be going from a mixture of group III and group I quadruplexes to a mixture of group I and group II quadruplexes.
Although for both temperatures there is no decrease in peak C, there is a decrease in intensity for peak B. The difference is larger for 20°C compared to that of the spectra at 10°C. The confirmations are maintained as a group I and group III mixture.

Figure 15. Change in CD spectrum over time at (a) 20°C and at (b) 10°C for sample at 5 µM in 0.01 M Tris buffer.
4.2 Crowding Experiments: PEG 600

In order to mimic intracellular crowding conditions, polyethylene glycol (PEG) 600 was added to our quadruplex folding and unfolding experiments. Spectra taken are represented in Figure 16 through Figure 19. Even at first glance, the CD spectra show prominent differences from one another, showing characteristics that may indicate different quadruplex groups, as well as the possibility of a hybrid mixture of two different types of quadruplexes.

Figure 16 through Figure 19 show the CD spectra for 10%, 20%, 30% and 40% of PEG 600. Figure 16 shows the sample with 10% (w/v) of the PEG 600. At the start of the experiment at 10°C, the spectrum shows a negative peak A, a positive peak B, and a positive peak C. As discussed prior, this pattern is consistent with a group II antiparallel conformation. As we increased the temperature to 90°C, we saw a significant decrease in peak C, which we also saw in the non-crowded DNA sample previously. While peak B just increases in intensity, peak C nearly disappears.

Larger changes were exhibited when the PEG 600 concentration was increased to 20% (w/v) (Figure 17). Here, the low-temperature spectra still resembles those of the 0% and 10% PEG, but with a subtle change: the starting curve has almost the same intensities for both peak B and peak C. Upon reaching 60°C, the high intensity for peak B is seen again, but this time without a significant bump for peak C. At 90°C, the whole spectrum loses its intense signal, and only has a residue of that of the lower temperatures: a negative peak A and a wide hump that lies between 260 and 290 nm. This subsequent decrease in intensity to the same shape at 70°C and 90°C indicate a denaturation of the entire structure.
At the highest concentrations of PEG, 30% (w/v) (Figure 18) and 40% (w/v), the temperature dependent spectra show a similar trend (Figure 19a). In these cases, the initial spectrum has a higher intensity peak B, with a diminished bump for peak C. Upon comparison to the data presented by Karsisiotis et al, the shape of this curve does not indicate a specific group. Instead, it can be speculated that if one overlays the group I and group II spectrum, our initial 10°C spectrum can be achieved. Thus for these crowding conditions, it can be concluded that the initial state of the sample contains a mixture of group I and group II conformations, which then switches to a completely parallel group I conformation upon heating from 10°C to 90°C. The minimized decrease in the peak B, and absence of peak C with an increase of PEG 600 to 40% PEG (w/v) sample indicates an increased stabilization of the DNA strand in the parallel quadruplex state.

Taking the previous time study into account, and noticing the importance of time on quadruplex systems, a repeat spectra of the 40% w/v PEG 600 sample was taken two months after the spectra taken in Figure 19. These newer spectra (Figure 20) also exhibit characteristics that lead us to believe that it is in a parallel conformation. There is a negative peak A, a high intensity peak B, and a shoulder for peak C exhibited only at the lower temperatures (10 to 30°C). Furthermore, just as before, we see a shift in peak intensities, with the decrease in peak B with increasing temperature being the most significant.

There is some difference in spectral shape between the earlier spectra, and the spectra taken 2 months later. The negative peak A is sharper and more negative in the new spectra than the first spectra taken, while the intensity at peak B decreased in return. However, the spectra still indicate the same conformation, leading us to conclude that the
sample in 40% (w/v) crowded conditions has a shorter equilibrium time than that of the sample under non-crowded solutions.
Figure 16. CD spectrum of sample in 5mM KCl, 0.01 M Tris buffer and 10% (w/v) PEG 600. Wavelength ranges from 230 to 320 nm, with the temperature being ramped up and down from 10°C to 90°C. At lower temperatures (10 to 60°C), the spectra show a negative peak A, a positive peak B, and a positive peak C indicating a possible mixture of antiparallel quadruplexes. At higher temperatures, the spectra have a negative peak A and a positive peak B, indicating a group I parallel conformation. With the temperature increases we see less clarity with the specific peaks at the different temperatures.
Figure 17. CD spectrum of sample in 5mM KCl, 0.01 M Tris buffer and 20% (w/v) PEG 600. Wavelength ranges from 230 to 320 nm, with the temperature being ramped up and down from 10°C to 90°C. At lower temperatures (10–60°C), the spectra show a negative peak A, a positive peak B, and a positive peak C indicating a possible mixture of antiparallel quadruplexes. At higher temperatures, the spectra have a negative peak A and a positive peak B, indicating a group I parallel conformation. With the temperature increases we see less clarity with the specific peaks at the different temperatures. Compared to that of the previous spectra at 0% and 10% PEG we see a definite increase for peak B, indicating a higher stabilization and concentration of the group I parallel conformation.
Figure 18. CD spectrum of sample in 5mM KCl, 0.01 M Tris buffer in 30% (w/v) PEG 600. Wavelength ranges from 230 to 320 nm, with the temperature being ramped up and down from 10°C to 90°C. The pattern is overall similar to the 20% spectra. However, we do see less of an intensity overall for peak C, indicating that the parallel conformation is even more favorable at this crowded state. Extreme temperatures (75 and 90°C) still show the tendency of overall denaturation as indicated by shallower peak B, which is maximized at 45°C.
Figure 19. CD spectrum of sample in 5mM KCl, 0.01 M Tris buffer and 40% (w/v) PEG 600. Wavelength ranges from 230 to 320 nm, with the temperature being ramped up and down from 10°C to 90°C. There is barely a hump for peak C compared to that of the previous spectra. All temperatures exhibit group I characteristics, with the hump disappearing as the run progressed. A denaturation at extreme temperatures still occurs.
Figure 20. CD spectrum of sample in 5mM KCl, 0.01 M Tris buffer and 40% (w/v) PEG 600 after 2 months of 15°C storage. Wavelength ranges from 230 to 320 nm, with the temperature being ramped up and down from 10°C to 90°C. With a negative peak A, a high intensity peak B, and a shoulder for peak C, we see that the quadruplex is still in the parallel conformation.
4.3 Crowding Experiments: PEG 8000

In order to mimic intracellular crowding conditions, PEG 8000 was also tested as the same weight to volume ratios: 10%, 20%, 30% and 40%. Although 10% and 20% show a similar pattern to PEG 600 in that it starts off with a possible combination of antiparallel conformations, at 30% and 40% the spectra are entirely different.

Figure 21 shows the CD spectrum of the sample in 5mM KCl, 0.01 M Tris and 10% of the PEG 8000. Through the entirety of the annealing and cooling process, the spectrum generally maintains the same shape with a negative peak A, a positive hump at B, and a high intensity peak C. The initial 10°C spectrum shows the largest intensities for all peaks and as the temperature increases to 60°C, peak B and peak C have intensities that are similar to one another. From 60°C upward to 90°C, the peaks become indistinguishable. The cooling of the system shows the return of the original shape of the spectrum, but the intensity of the final 10°C spectrum is not even half of that of the original 10°C spectrum. We can also compare this to that of the spectrum with no crowding agent, and can note how they are similar, except for the fact that the crowded system lacks the recovery of the signal during the cooling process. Although it is tempting to conclude that this shows that the interactions of the system are non-reversible and is thus it is not a two state system, our time study mentioned previously must also be taken into account. We must conduct additional studies to insure that the system has been able to reach equilibrium.

Although the 20% (w/v) PEG 8000 spectrum could not be corrected for its baseline due to instrumental issues the overall pattern in its shape alteration with change in temperature with this higher concentration is apparent (Figure 22). The initial state is
still the same as that of the 10% (w/v) PEG 8000 spectrum at 10°C with a negative peak A, a positive peak B and a high intensity peak C. With the temperature increase from 10°C to 45°C, the height of peak C decreases until it is similar to that of peak B. Subsequently with the increase of temperature from 45°C to 90°C, both peaks become indistinguishable from one another to form a wide peak that spans from 250 nm to 340 nm. The annealing process exhibits a different pattern; although the peaks became indistinguishable at around 75°C, the cooling doesn’t resolve these two peaks until the 30°C mark.

A large change in pattern occurs at the 30% (w/v) PEG 8000 sample, as seen in Figure 23. Instead of showing the distinct peaks A, B and C that we have used to compare to other groups, such as Karsisiotis et al, this sample has a novel shape. At all temperatures the spectra exhibit the same basic form: a downward dip at around 240 nm, reminiscent of peak A, but not quite present at all temperatures, and a small, wide peak at around 270 nm. One would be tempted to assume that the positive peak and the seeming negative peak are synonymous to those of a group I complex. However, here the positive peak is at around 280 nm, and there is no distinct negative peak A, only a large downward sloping line. Furthermore, although we ramped the sample temperature from low to high and back again as we did with our previous samples, the intensity of the peak barely changed as the experiment was conducted, and there are no clear patterns to associate the change in signal to temperature. The 40% (w/v) PEG 8000 sample, as seen in Figure 24 (top), also exhibits a similar phenomenon.

Notably, this set of spectra of 30% (w/v) and 40% (w/v) PEG 8000 is almost the same as that of the cytosine-containing complementary strand, as seen in a spectra
comparison shown in Figure 24 (bottom). The curves are similar in shape and trend, although the intensity of the spectra obtained at high PEG concentrations is much lower than that of the spectra for the complementary strand. Thus, it appears that the quadruplex conformation is severely destabilized by high crowding conditions, which is a conclusion not seen in most of the literature and will be addressed in the discussion,
**Figure 21. CD spectrum of sample in 5mM KCl, 0.01 M Tris buffer and 10% (w/v) PEG 8000.** Wavelength ranges from 230 to 320 nm, with the temperature being ramped up and down from 10°C to 90°C. At lower temperatures (10 to 60°C), the spectra show a negative peak A, a positive peak B, and a positive peak C indicating a possible mixture of antiparallel quadruplexes. At higher temperatures, the spectra have a negative peak A and a positive peak B, indicating the possibility of a group I parallel conformation. This spectrum is very similar to that of the non-crowded sample.
Figure 22. CD spectrum of sample in 5mM KCl, 0.01 M Tris buffer and 20% (w/v) PEG 8000. Wavelength ranges from 230 to 320 nm, with the temperature being ramped up and down from 10°C to 90°C. Although the baseline was unable to be corrected in this spectrum, the pattern is still apparent. Initially at 10°C, the CD spectrum exhibits the same shape as the previous spectra have for the initial state: a negative peak A, a positive peak B, and a larger intensity for peak C. With the heating of the system from 10°C to 45°C, peak C decreases until its height is comparable to that of peak B. From 45°C to 90°C both peaks decrease to a large and wide peak that spans between 250 and 340 nm. Cooling from 90°C to 60°C, only shows an increase in intensity of the large, wide peak that was present at 90°C. From there down to 10°C, peak B and C become more distinct from one another, to have a final spectral shape of a negative peak A, a positive peak B, and a positive peak C at 10°C. Again the signal is not comparable to that of the original spectrum and thus this process is non-reversible and not two state.
Figure 23. CD spectrum of sample in 5mM KCl, 0.01 M Tris buffer and 30% (w/v) PEG 8000. Wavelength ranges from 230 to 320 nm, with the temperature being ramped up and down from 10°C to 90°C. The shape of this curve is distinct from those of smaller PEG 8000 concentrations. There is a negative dip at 240 nm that perhaps is synonymous with that of peak A we have seen in previous spectra, but that is uncertain. Furthermore, instead of having a peak at 260 nm and at 290 nm, there is large flatter peak at around 280 nm instead. Even with such large changes in temperature, the overall spectra neither have significant changes in shape or intensity. It is also noteworthy that there is no clear pattern with the slight deviations in signal for different temperatures.
Figure 24. Comparing CD spectra of 40% (w/v) PEG sample (top) with complement spectra (bottom) Looking exclusively at shape, both spectra exhibit a peak at around 280 nm, although the two spectra differ in the 230 to 260 nm range. The semi dip at 240 nm in the 40% PEG spectra can be an indication of lingering quadruplex formations. This can also be determined by the fact that this is disappearing as the temperature increases, assuming that heat denaturation still occurs at this highly crowded state. The bottom spectra was taken by Sally Shepardson-Shepardson-Fungairino, at room temperature using the complement sequence at a 5µM DNA concentration in 0.01M Tris buffer.
4.4 Crowding Experiments: Ethylene Glycol

After exploring quadruplex formation in the presence of the larger polymeric crowding agents PEG 600 and 8000, we tested ethylene glycol (EG) at the same concentrations. Assuming that the larger agents had excluded volume effect, we aimed to use EG as a measure of a different kind of crowding, i.e. the effects of water activity on quadruplex formation. The results can be seen Figure 25 through Figure 28.

Figure 25 exhibits the CD spectra of our sample with 10% (w/v) ethylene glycol, with 5 mM KCl in 0.01 M Tris buffer. As before, spectra were taken from 10°C to a maximum of 90°C (heating), and then taken from 90°C back down to 10°C (annealing). Similar to the spectra obtained with no or low concentrations of polymeric crowding agents, we see the same shape here with a negative peak A, a positive peak B, and a larger peak C at our initial 10°C spectrum. Then, while the temperature increases from 10°C to 45°C, peak B remains unchanged, while peak C significantly decreased. Furthermore, when the temperature of the system increases to 60°C, we see a sudden decrease in peak B and C, and become comparable in size. From 60 to 90°C, those two peaks became indistinguishable and form a wide peak between 250 nm and 300 nm. This phenomenon was also seen for higher temperatures in most of the other experiments we had conducted. Surprisingly, in this case, we almost see reversibility in our system, as apparent by the cooling curves on the CD spectrum: they follow the opposite pattern of that of the annealing curves.

A similar pattern can be seen in the 20% (w/v) EG sample (Figure 26), as well as the 30% (w/v) EG (Figure 27) and the 40% (w/v) EG (Figure 28). However, though they have the same patterns, it is important to realize the differences that these changes in
concentration of crowding agents bring. One, the height of peak C decreases as the crowding agent concentration is increased. Second, although you see this change in peak C, at 60°C for all concentrations, peak B and peak C are both at the same relative ellipticities.
Figure 25. CD spectrum of sample in 5mM KCl, 0.01 M Tris buffer and 10% (w/v) EG. Wavelength ranges from 230 to 320 nm, with the temperature being ramped up and down from 10°C to 90°C. Initially, as we often see in the initial of our experiments, there is a negative peak A, a positive peak B, and a larger peak C. As the temperature is increased from 10°C to 45°C, peak B is maintained, but the signal for peak C decreases. At 60°C, there is a sudden decrease in both of these peaks, with both signals dropping to about the same ellipticity. The subsequent higher temperature scans involve peak B and C becoming indistinguishable from one another, reducing to a small and wide between 250 nm and 300 nm, as we’ve seen before. As the sample goes through the cooling process, it follows the opposite pattern as that of the annealing of the sample, indicating more of a reversibility, peak A remains relatively unchanged as the temperature is increased and decreased.
Figure 26. CD spectrum of sample in 5mM KCl, 0.01 M Tris buffer and 20% (w/v) EG. Wavelength ranges from 230 to 320 nm, with the temperature being ramped up and down from 10°C to 90°C. The spectrum is very similar to that of the 10%(w/v) EG experiment. The basic shape and the pattern of change exhibited over various temperatures are very similar. It is important to note the decrease in signal for the lower temperature peaks, while the higher temperatures (60°C and upward) are at the same relative intensity as that of the 10% EG sample.
Figure 27. CD spectrum of sample in 5mM KCl, 0.01 M Tris buffer and 30% (w/v) EG. Wavelength ranges from 230 to 320 nm, with the temperature being ramped up and down from 10°C to 90°C. The spectrum is very similar to that of the 10% and 20% (w/v) EG experiment. The basic shape and the pattern of change exhibited over various temperatures are very similar. It is important to note the decrease in signal for the lower temperature peaks, while the higher temperatures (60°C and upward) are at the same relative intensity as that of the 10% and 20% (w/v) EG experiment.
Figure 28. CD spectrum of sample in 5mM KCl, 0.01 M Tris buffer and 40% (w/v) EG. Wavelength ranges from 230 to 320 nm, with the temperature being ramped up and down from 10°C to 90°C. The spectrum is very similar to that of the 10%, 20%, and 30% (w/v) EG experiment. The basic shape and the pattern of change exhibited over various temperatures are very similar. It is important to note the decrease in signal for the lower temperature peaks, while the higher temperatures (60°C and upward) are at the same relative intensity as that of the 10% and 20% (w/v) EG experiment.
5. Discussion

5.1 Time vs. Temperature

To understand fully the time dependence and temperature dependence of the quadruplex CD spectra, we must first understand the factors that go into the specific conformation and stability of a formed G-quadruplex. These factors include presence and identity of the cation, the DNA sequence, and the environment (which will be discussed in the following sections on crowding studies). We must also be aware, since transcription and translation can occur at a rate of $10^{-2}$ sec$^{-1}$, that the unfolding and folding kinetics of a quadruplex must be taken into consideration relative to this value.

The cation coordinates itself inside the guanine tetrad, and situates itself inside of a cavity formed by eight O$_6$ atoms, four in the tetrad above and four below. Due to the relative sizes of the cavity and the cation, K$^+$ better stabilizes the structure than Na$^+$. Additionally, quadruplex stability is dependent on the hydration of the system, which is intricately linked to the cation, Na$^+$ vs. K$^+$ that shields the negatively charged backbones in solution. Hud et al. examined the competition between these two monocations for coordinating G-quadruplexes. Although the researchers noticed a $1.7 \pm 0.15$ kcal•mol$^{-1}$ difference between the folding free energy $\Delta G^\circ$ between the two ions, the difference in energy of hydration for the two ion quadruplex conformations was found to be around 17.6 kcal•mol$^{-1}$ for the system, around 10 fold larger. Therefore, in order to facilitate quadruplex formation, our group decided to conduct our experiments with potassium monocations.

The sequence of the strand also changes the most stable conformation of the quadruplex. Olsen et al. conducted a comprehensive study on the unfolding of G-
quadruplexes using CD, UV spectroscopy, and Differential Scanning Calorimetry (DSC) to understand the thermodynamic characteristics of the systems as well as the contributions of water and ion to the overall stabilities. They tested four different biologically relevant sequences: $d(G_2T_2G_2TGTG_2T_2G_2)$ (renamed G2), $d(G_3T_2G_3TGTG_3T_2G_3)$ (renamed G3) are both DNA aptamers; $d(TG_4AG_3TG_4AG_3TG_4A_2G_2)$ or NHE-III which is in the promoter region of c-MYC oncogene; and a human telomeric sequence $d(AG_3T_2AG_3T_2AG_3T_2AG_3T_2)$ (renamed 22GG). With these different sequences, different conformational structures were exhibited: G2 was all antiparallel, NHE-III and 22GG showed all parallel conformations, while G3 showed a unique parallel and antiparallel mixture ("basket" and "chair"). Therefore the study concluded that 3 or more guanines in a row can cause aggregation, and many folded states are possible.  

Olsen et al. also found that their mixed system with the G3 strand was sensitive to changes in temperature and sample preparation. They used a “short equilibrium time” experiment (simply dissolving of the DNA in buffer) and a “long equilibration time” experiment (slow annealing followed by constant sample temperature), each of which exhibited different spectra. Their results, shown in Figure 29, are reminiscent of our own.

**Figure 29. CD spectra of G3 in 10 mM Cs-HEPES buffer, 100 mM K+ at pH 7.5 at 20°C by Olsen et al.** The inverted triangles depict the short temperature equilibrium time, while the triangles represent the long temperature equilibration time experiment. As you can see, the long equilibrium exhibits a higher peak B compared to that of peak C, which is a clear indication of a mixed conformation. The difference in spectra exhibit the high sensitivity of the system.
(Figure 14). As we found in our study, our sample was sensitive to temperature and exhibited a possible mixed conformation. Indeed, even the shapes of the spectra are similar. The long equilibrium is exhibited in the 10°C spectra for day 0, where there is a high intensity peak B and a shoulder for peak C. The subsequent scans may be indicating the short equilibrium, although with our higher intensity peak C, this may be indicating a higher concentration of antiparallel quadruplexes compared to that of the Olsen study.

Olsen et al. also carefully investigated the temperature dependence of folding, taking their temperatures as low as 20°C and going up to 105°C at a rate of 0.9°C/min. As their results show in Figure 30, when the temperature increases, the denaturation is apparent, with the decrease of peak B to almost 0 millidegrees as the temperature is increased, in addition to the decrease of peak C. This basic pattern is seen in our Figure 13 as well.

When taking kinetics into consideration, Zhao et al. measured rate constants on the order of $10^{-3}$ s$^{-1}$ for quadruplex folding in the presence of potassium ions using SPR. In contrast, the studies for this project have exhibited changes over a prolonged period of time for quadruplex formation. It would be supposed that both the sequence used in the Zhao study, (TTAGGG)$_4$, and the study here would be capable of forming an

---

**Figure 30. CD spectra of G3 sequence in 10mM Cs-HEPES buffer, 100 mM K$^+$ at pH 7.5 from 20°C to 105°C by Olsen et al.** The initial spectrum shows a negative peak A, a high ellipticity peak B, and a lower peak. This is similar to that of our Figure 14 Day 0 study in terms of conformation and destabilization over temperature.
intermolecular quadruplex. Differences may be occurring due to the relative concentration of salt, which was 150 mM K\(^+\), while it was 5 mM K\(^+\) in our study. The biggest flaw in this cited study may be that they assumed a two state system, while we are open to the possibility of a change over a prolonged period of time. The assumption would take the traditional model of a two-state melting curve as a basis, disregarding the slope that would occur on the two seemingly flat lines which sandwich the steep portion of the melting curve. This multi-state system may have possible connections to the difference in salt concentrations (5 mM vs. 150 mM), sequence and the competition and relative stabilities of multi-stranded and single-stranded quadruplexes.

Instead, to understand the isothermal conformational changes that we found with our spectra measured over several days (Figure 15), we can refer to Rehm et al.’s article in Chemical Science, in which they explored the sequence d[(G\(_4\)CT)\(_3\)G\(_4\)], which is associated
with the human pathogen *Treponema pallidium*. CD spectra for these studies are shown in Figure 31 and are all aimed to find the isothermal folding constants. Subplot a shows the CD spectrum of the sample that was denatured and slow-cooled without any $K^+$ present. The before line (red) shows the CD spectrum for when the strand is denatured, the blue line indicates the addition of 1 mM of KCl, which was then observed for 75 hours (subsequent CD spectra are in blue and green). With the spectra being taken at 20°C, they observed that the strand folded into an antiparallel spectrum within the 75-hour time period. Subplot b was denatured and slow-cooled with 1 mM of $K^+$ present. The final concentration of $K^+$ was then increased to 500 mM and observed over 75 hours. Even with a high concentration of salt added a high activation energy for unfolding and refolding prevents a change in the conformation of the sample. Thus, depending on the stabilizing effects of the surrounding environment, you will come across instances where the structure is stabilized enough that it will stay in its original conformation.

Moreover, the system here has come to an equilibrium after 75 hours and is highly stabilized to the point where the addition of more salt does not induce any changes. To compare this with our results, we must conduct our isothermal experiment over the 75 hours to see if we end up with a higher ratio of group III antiparallel quadruplexes as this group observed.

Interestingly, various laboratories have sometimes encountered conflicting results and come to different conclusions as they aim to determine the kinetics behind quadruplex formation. In his paper “Human telomeric G-quadruplex: thermodynamic and kinetic studies of telomeric quadruplex stability” Johnathan B. Chaires discusses the inconsistencies between all of these and brings up several factors that may be
contributing. In terms of the energetics of unfolding human telomeric quadruplexes, \( T_m \) values range from 63 to 81.8°C, with an enthalpy range of 49 to 77.5 kcal mol\(^{-1}\) in 100 mM K\(^+\) concentration solutions. All data were obtained on similar sequences as well as similar solution concentrations. Although to some degree the discrepancies can be attributed to the difference in fluorescence labels, annealing procedures, or difference in sequences, even studies that use the same sequences seem to have varied results from one another.\(^{29}\)

Chaires effectively lays out some possible sources of difference between studies of quadruplex folding. The first is the discrepancy in finding the pre- and post-transition baselines. These can be sloping, which may arise from other physical phenomena, such as the expansion of solvent or additional reactions. In our studies, the 0.01 M Tris buffer blank exhibits temperature dependence, as seen in Figure 32. We even see a curve for the PEG blank, which exhibits a dipped curve as seen in Figure 33. Furthermore, the mere assumption that denaturation is only a two-state process can be deterring from the true and best results. However, this can be avoided by overlaying the obtained denaturation curves from two or more methods of data collecting. If the two are superimposable, that implies that the mechanism is indeed two-state. In addition, the researcher could also use

### Table 2. Energetics of human telomere quadruplex unfolding

*Adapted from Chaires, J.B. (2010)*\(^{29}\)

<table>
<thead>
<tr>
<th>Sequence 5' --&gt; 3'</th>
<th>(<a href="%5Ctext%7BmM%7D">K^+</a>)</th>
<th>(T_m(\text{°C}))</th>
<th>(\Delta H(\text{kcal/mol}))</th>
<th>(\Delta S(\text{cal/mol/K}))</th>
<th>(\Delta G_{310}(\text{kcal/mol}))</th>
</tr>
</thead>
<tbody>
<tr>
<td>(TTAGGG)(_4)</td>
<td>70</td>
<td>63</td>
<td>49</td>
<td>147</td>
<td>3.4</td>
</tr>
<tr>
<td>AGGG(TTAGGG)(_3)</td>
<td>100</td>
<td>63</td>
<td>57</td>
<td>169</td>
<td>4.6</td>
</tr>
<tr>
<td>GGG(TTAGGG)(_3)</td>
<td>100</td>
<td>65</td>
<td>60.5</td>
<td>179</td>
<td>5</td>
</tr>
<tr>
<td>GGG(TTAGGG)(_3)</td>
<td>100</td>
<td>69.3</td>
<td>77.5</td>
<td>202</td>
<td>14.8</td>
</tr>
<tr>
<td>TGGG(TTAGGG)(_3)</td>
<td>100</td>
<td>81.8</td>
<td>66.2</td>
<td>186.5</td>
<td>8.4</td>
</tr>
<tr>
<td>AGGG(TTAGGG)(_3)</td>
<td>100</td>
<td>66.1</td>
<td>34.4</td>
<td>101.4</td>
<td>3</td>
</tr>
</tbody>
</table>
a dual-wavelength parametric test, where the two data sets are plotted against one another.

Therefore, to confirm the fact that the system is not at equilibrium and irreversible, we conducted a melting study. A sample at 5 µM DNA concentration was put into the CD spectrometer, going through the same temperature sequence as that of the original temperature and time study. We took CD measurements for every temperature point (10, 20, 30, 40, 60, 75 and 90°C) and ramped it back down at the same rate and speed as before. The CD spectra of these melting curves at 265 nm and 290 nm can be seen in Figure 34 top and bottom, respectively.

The 265 nm curves show a distinct irreversibility and the non-two state properties of this system under these experimental conditions. The ramp up shows some characteristics that are reminiscent of a traditional two-state melting curve: a flatter region that would have extended from a lower temperature to 20°C, a steeper curve ranging from 20°C to 45°C, and then a flatter curve from then and upwards to 90°C. The cooling curve has no such shape and is not superimposable onto the melting curve. Although the curve seems to be along the same trend from 90°C to 45°C, there is a distinct variation from 45°C down to 10°C. Instead of curving upwards as the annealing curve had, the cooling curve then dips down even further, hitting a minimum in CD measurement at around 20°C.

The 290 nm curves aren't as apparent in this effect, but are still non-superimposable. However, they both characteristically mimic the shape of the traditional two state curve: there is a shallower curve from 10 to 30°C, a steeper curve from 30°C to 50°C, which then levels out to 90°C. Both the annealing and cooling curves have the
transition temperatures for the slopes of these curves, but vary in signal size from 10 to 40°C, with the biggest gap at the lowest temperature.
Figure 32. CD spectrum of 0.01 M Tris buffer blank. Blank spectra were taken in parallel with all sample runs under the same conditions where the temperature was heated from 10°C to 90°C and cooled back down. As one can see, the blank exhibits temperature dependency. The curve shifts upward, but maintains the shape as the temperature is increased.
Figure 33. **CD spectrum of 40% PEG 600 solution.** Blank spectra were taken in parallel with all sample runs under the same conditions where the temperature was heated from 10°C to 90°C and cooled back down. As one can see, the blank exhibits temperature dependency. The curve shifts upward, but maintains the shape as the temperature is increased. The PEG exhibits a characteristic dip at 240 nm, which levels out at higher wavelengths.
Figure 34. CD melting curves for 5µM DNA sample in 0.01 Tris Buffer at 265 nm (top) and at 290 nm (bottom). The top plot shows that the 260 nm CD measurement decreases as the temperature is increased and resembles a traditional two state melting curves with a steeper slope in between two, almost flat lines on the curve. However, the cooling and heating do not have the same shape. Although they may be similar from 50° to 90°C, the lower temperatures lead us to the irreversibility and non-two state characteristic of this system at the speed we are taking the spectra. The bottom plot for the 290 nm is not quite as apparent for the nature non-reversible notion of the system. The non-superimposable nature of the two plots are also an indication of this property.
5.2 Crowding Experiments

As discussed in the introduction, the various crowding agents that we are using vary in molecular weight, and thus should have different effects of the formation of quadruplexes in our experiments. Since the molecular composition of the monomer units of the polymer used in this series are the same, it is necessary to look at the gyration diameters. As the size of the agent greatly affects a biological system, as seen in a study done by Rajiv Bhat et al. at Brandeis University, citing that the radius of gyration is related to the molecular weight by the equation:

\[ R_g = aM^{0.55} \]  

with \( a \) as an experimentally found constant, and \( M \) the molecular weight.\(^{30}\) PEG 8000, being a large molecule, should be creating area that is inaccessible to other molecules and consequently increasing the viscosity of the solution. This volume exclusion may increase the thermodynamic activity and therefore may increase the association rate.\(^{7}\)

On the other hand, the smaller crowding agent EG is not large enough to act as obstacles and will not induce an excluded volume effect; instead, the water activity is decreased which creates osmotic pressure. It is especially crucial here to note since most nucleic acid reactions involve either the release or the uptake of water.\(^{7}\) There are 11 to 12 hydration sites on a nucleotide, and these interact with water to create a hydration shell. The creation of these shells has a strong role in determining the structure and stability of the building blocks of DNA.\(^{22}\)

Yong Xue et al., also conducted a study on parallel-stranded intramolecular G-quadruplexes in K\(^+\) solutions under molecular crowding conditions using PEG 200 for a telomeric DNA sequence, \( G_3(T_2AG_3)_3 \). They achieved the spectra as seen in Figure 35A.
With a K\(^+\) concentration held at 150 mM K\(^+\) to mimic intracellular conditions, PEG 200 was added at concentrations of 10, 20, 30, and 40\% (w/v). With no PEG present, the spectrum exhibits a negative peak A, a slight hump for peak B, and a positive peak C, which is similar to that of our spectrum as seen in Figure 13. Then, as the crowding agent concentration is increased, the hump for peak C gradually forms into a well-formed peak, while the peak B disappears, which we also see by comparing Figure 16 through Figure 19 for PEG 600. This indicates a conformational switch for both the Xue study as well as ours, from an antiparallel conformation to a parallel one. It is also noteworthy that the same study, but with Na\(^+\) as the cation instead (Figure 35b), the crowding agent does not affect the stability of the antiparallel complex formed under these conditions.\(^1\)
Figure 35. CD spectra of G_3(T_2AG)_3 in 150 mM K^+ solution (A) and 150mM Na^+ solution with various concentrations of PEG. Panel a for the potassium solution shows a clear change in conformation based upon the differences in peak location and heights as the concentration of PEG was increased. The PEG induced a conformational change from an antiparallel complex to that of a parallel one. Panel b is important to note for the difference in Na^+ and K^+ and how the identity of the ion still has a key role in structure determination.\(^1\)
As discussed earlier, Buscaglia et al. conducted a study using several different types of crowding agents: ethylene glycol, diEG, triEG, PEG 200, PEG 400, PEG 600, PEG 1000, PEG 1500, PEG 3350, PEG 8000 and PEG 10000 with their DNA sequence of choice, 5’-d(AGGG(TTAGGG)₃)-3’. Figure 36A shows all of these crowding agents at the highest concentration used in this study, 42.5% (v/v), and the group suggests that all of these crowding agents exhibit a change in conformation from an antiparallel conformation to that of a parallel one. In spite of these claims, our data disagrees with this study, as we saw a clear difference in trends between crowding agents.

However, perhaps these differences provide a clear indication for a need in a systemized method. For Buscaglia et al., for their co-solvent titrations, samples were prepared at 4 µM strand concentration and annealed slowly. The annealing method involved placing the sample in 1 L of boiling water for 20 minutes and the bringing it to room temperature by slow cooling. These were then allowed to incubate in a water bath overnight before the spectra were taken.
Indeed, Karimata et al. present a spectrum that looks quite distinct from that of the study done by Buscaglia et al.; they used two different sequences: Tetrahymena (intra-Tet, d[T2G(G3T2G)3G]) and intermolecular human (intra-Hum, d[G3T2A]3G3). There are many factors that may be attributed to the difference in CD spectra between the Karimata et al. study and the Buscaglia et al. study. For one, it may be the identity of the salt (10 mM sodium phosphate buffer (pH 7.0), 100 mM NaCl, and 1mM Na2EDTA). Another difference is the rate at which the CD spectra were taken; here they claim that the Tm value is not affected by the heating rate if it is between 0.1 to 0.5 C° min⁻¹. Furthermore, Karimata et al.’s spectra were obtained at 4°C, as opposed to Buscaglia et al.’s 25°C experimental condition, post slow annealing.⁴

Furthermore, Karimata et al.’s two sequences, intra-Tet and intra-Hum also exhibited different structural characteristics under the very same conditions, an indication of the importance of the identity of the DNA strand. The crowding agents, EG, di(EG), tri(EG), PEG 200 and PEG 2000 were pipetted into the sample with a 40 % (w/v) concentration. Intra-Hum exhibited no change in shape in the spectra with a change in crowding agent: there’s a positive peak A, a negative peak B, and a positive peak C indicating a group III antiparallel complex (Figure 37b). The group points out that this demonstrates that intra-Hum does undergo any type of antiparallel to parallel conformational transition. On the other hand, the intra-Hum crowded samples show large deviations from the spectra of the non-crowded sample (Figure 37a). With no cosolute the CD spectra shows a negative peak A, a positive B and a positive peak C, indicating a mixture of a parallel and antiparallel complexes. An increase in molecular
cosolute weight leads to an increase in the intensity of peak B, indicating the conformational transition from antiparallel structures to parallel ones.\textsuperscript{4}

The group goes on to claim that in general, all molecular crowding conditions should initiate the association of monomers into multi-strand complexes, which may lead to structures such as a multi-stranded G-wire.\textsuperscript{4} A G-wire is a one-dimensional supramolecular structure, which consists of stacked guanine tetrads.\textsuperscript{31} Karimata et al. claim that this must occur since the parallel structure exhibited by the intra-Tet must be multi-stranded. With a UV-melting study, with an increase in $T_m$, increased from 60.8 to 65.0$^\circ$C, to indicate an increased stability. Thus the group concluded that molecular crowding agents induce structural change and stability of the quadruplex structures.\textsuperscript{4} This may also be the case here, as the long equilibration time may be indicating the possibility of multiple strands coming together to form the group I parallel quadruplexes at high crowding agent concentrations. This can also be determined by the fact that the criteria
for forming a quadruplex cannot be met to construct a parallel conformation without using multiple strands.

Daisuke Miyoshi et al. present several studies that used various concentrations of the crowding agents PEG 200, PEG 400, and EG. As the data presented here indicates, with the increase in melting temperature and the decrease in $\Delta G^\circ$, it confirms the theory that crowding stabilizes the quadruplex. By crowding, if you recall the thermodynamic equation $\Delta G = \Delta H - T\Delta S$, the formation of the quadruplex enhanced by the enthalpic contribution was able to overcome the negative entropic contribution.\(^{32}\) Moreover, in conjunction with this being true, there are also other effects, such as hydration, that may be contributing to quadruplex conformation and stability.

We then must also consider the hydration of the quadruplex, which Miyoshi et al. gave compiled into a table as seen in Table 4. For different structures, osmotic stress techniques have found that there are a different number of water molecules bound, a value that be represented as $-\Delta n_w$. This can be calculated using the following equation

$$-\Delta n_w = n_{w,\text{folded}} - n_{w,\text{unfolded}}$$

(7)

where $n_{w,\text{folded}}$ and $n_{w,\text{unfolded}}$ represent the number of water molecules bound to the folded state and the unfolded states.\(^{32}\) A number of factors can go into the uptake or release of water molecules. Olsen et al. list a few possibilities that rationalize their numbers:

1. There is a release of water molecules when there is a structural transition from a random coil to a quadruplex conformation.

2. The quadruplex has a higher charge density parameter than a random coil, and thus has a higher number of water molecules surrounding it.
3. There is a release of water when K$^+$ binds to the G-quadruplex core, which will result in a smaller difference in water binding.$^5$

Even though we see a large range in numbers for $\Delta n_w$ (from 0.14 to 8.5), this data set still allows us to conclude that all G-quadruplex formations result in dehydration, which has also lead to their increased stability under crowding conditions. This is one clue that will lead us one step closer to understanding quadruplex formations in vivo.
**Table 3. G-quadruplex stabilities under molecular crowding conditions.** Adapted from Nakano et al. (2012)\(^7\)

<table>
<thead>
<tr>
<th>Sequence</th>
<th>Cosolute</th>
<th>(\Delta T_m) (°C)</th>
<th>(\Delta \Delta G^\circ) (kcal/mol)</th>
<th>(\Delta \Delta H^\circ) (kcal/mol)</th>
<th>(\Delta \Delta S^\circ) (kcal/mol)</th>
</tr>
</thead>
<tbody>
<tr>
<td>G(_3)T(_3)G(_3)T(_3)G(_3)</td>
<td>30% EG</td>
<td>12</td>
<td>-3.5</td>
<td>-16</td>
<td>-12</td>
</tr>
<tr>
<td>G(_3)T(_3)G(_3)T(_3)G(_3)</td>
<td>30% EG</td>
<td>15</td>
<td>-2.4</td>
<td>-6.4</td>
<td>-3.8</td>
</tr>
<tr>
<td>(G(_3)T(_2)A(_2))G(_3)</td>
<td>0.75 M PEG 200</td>
<td>11</td>
<td>-4.6</td>
<td>39</td>
<td>43</td>
</tr>
<tr>
<td>(G(_3)T(_2)A(_2))G(_3)</td>
<td>0.75 M PEG 400</td>
<td>14</td>
<td>-2.1</td>
<td>69</td>
<td>71</td>
</tr>
<tr>
<td>bcl-2 DNA</td>
<td>40% PEG</td>
<td>26</td>
<td>-4.5</td>
<td>-12</td>
<td>-7.6</td>
</tr>
</tbody>
</table>

**Table 4. Hydration of G-quadruplexes.** Adapted from Miyoshi. et al (2010)\(^32\)

<table>
<thead>
<tr>
<th>Sequence</th>
<th>Structure</th>
<th>[K+] (mM)</th>
<th>(\Delta n_w)</th>
</tr>
</thead>
<tbody>
<tr>
<td>G(_2)T(_2)G(_2)TGTG(_2)T(_2)G(_2)</td>
<td>Antiparallel (chair)</td>
<td>100</td>
<td>4.5</td>
</tr>
<tr>
<td>G(_2)T(_2)G(_2)TGTG(_2)T(_2)G(_2)</td>
<td>Antiparallel (chair)</td>
<td>100</td>
<td>0.53</td>
</tr>
<tr>
<td>G(_3)T(_2)G(_3)TGTG(_3)T(_2)G(_3)</td>
<td>Antiparallel (chair)</td>
<td>100</td>
<td>1.1</td>
</tr>
<tr>
<td>TG(_4)AG(_4)TG(_4)AG(_4)TG(_4)A(_2)G(_2)</td>
<td>Parallel</td>
<td>100</td>
<td>1.5</td>
</tr>
<tr>
<td>AG(_5)T(_2)AG(_5)T(_2)AG(_5)T(_2)AG(_5)</td>
<td>Antiparallel</td>
<td>100</td>
<td>0.14</td>
</tr>
<tr>
<td>G(_2)T(_2)G(_2)TTTG(_2)T(_2)G(_2)</td>
<td>Antiparallel</td>
<td>100</td>
<td>1.4</td>
</tr>
<tr>
<td>G(_3)T(_3)G(_3)T(_3)G(_3)T(_3)G(_3)</td>
<td>Parallel</td>
<td>100</td>
<td>0.27</td>
</tr>
<tr>
<td>G(_3)T(_3)G(_3)T(_3)G(_3)T(_3)G(_3)</td>
<td>Mixed</td>
<td>100</td>
<td>0.14</td>
</tr>
<tr>
<td>A(G(_3)T(_2)A(_2))G(_3)T(_2)</td>
<td>Mixed</td>
<td>100</td>
<td>0.17</td>
</tr>
<tr>
<td>bcl-2 DNA</td>
<td>Parallel</td>
<td>5</td>
<td>8.5</td>
</tr>
</tbody>
</table>
6. Conclusion

By studying in vitro formation of DNA G-quadruplexes in the *Bdellovibrio bacteriovorus* genome, by the sequence 5'-GGGTGGGTCAGGGGCAGGG-3', we discovered clear conformational differences that occur as a function of folding time, temperature and crowding environment.

With the time experiments, we saw a dependence on folding time for the quadruplex formations on the order of hours to days. Under non-crowding conditions, the sequence formed different mixes of group II and group III quadruplex conformations when spectra were obtained over the span of several days. Constant temperature spectra measured at 10°C and 20°C over the course of 20 hours also showed some denaturation, indicated by the decrease in peak B over the time span. However, during a singular run, the spectra showed denaturation as the temperature increased; the difference in relative intensity of peak B and peak C also indicate a possible change in the structural conformation ratio.

The crowding experiments were conducted with PEG 600, PEG 8000 and EG, with each showing different effects with respect to concentration of the crowding agent. With PEG 600, a clear conformational shift occurred from an antiparallel quadruplex mixture to a parallel conformation. With PEG 8000, the pattern was not quite as clear; it began off with a similar pattern in that there is a shift from an antiparallel mixture to a parallel one. However, once the concentration reaches 30 and 40% (w/v) for PEG 8000, the spectra resemble that of the complement strand. On the other hand, EG shows a completely different pattern: there is no conformational shift with increasing
concentration of crowding agent and the quadruplexes remain in a mixed antiparallel state.

One additional variable that should be introduced in our study is the DNA sequence itself. A fellow researcher in the lab, Sally Shepardson-Fungairiño, did a study of a several sequences from the same genome, which yielded different conformations for different sequence; this indicates an importance in our sequence choice. To systematically test this, it may be best to interchange non-guanine nucleotides, or test the length of loops to see whether these may affect the conformation we get under the various variables we tested in this study.

To further support our claims, future studies will need to involve detailed attention to the role of kinetics in our experiments. With the long folding times observed here, we cannot yet evaluate which structures are most likely to occur in vivo. One of the next steps involves using differential scanning calorimetry to measure the exact constants associated with the folding, unfolding or conformational shifts of the systems that we are observing. Additionally, taking an initial isothermal CD spectrum over 3 days or more with our sample will indicate the stability of the quadruplexes in its initial conditions. With our discovery that the folding of these sequences cannot be described by a purely two state model, alternate ways of qualitative analysis must be developed; these methods, with controlled variables will ultimately lead to the possibility of more quantitative analysis.

Moreover, it is clear that the folding and storage history of our samples must also be controlled. The number of times the DNA sample has been heated and annealed may contribute to the possibility of differentiating spectra for that of the same sample. We
should also address the issue with controlling the DNA concentration of our samples. Here, the crowding agents have been titrated in; thus, with the addition of crowding agents, the DNA concentration decreases as the crowding agent concentration increases.

By controlling all of these new found variables and careful reconsiderations of data analysis, both qualitative and quantitative, we will be able to come one step closer to understanding the conditions of DNA G-quadruplex formation as well as the characteristic forms these secondary structures will take on. By diversifying our crowding agent library to include such molecules as proteins, we will be able to apply our knowledge to quadruplex formations to that of *in vivo*.
7. References


