Spring 5-2-2021

Chemical and Co-Solute Effects of Polyethylene Glycol on i-Motif Formation

Lindsey Rutherford

CHEMICAL AND CO-SOLUTE EFFECTS OF POLYETHYLENE GLYCOL ON I-MOTIF FORMATION

By
Lindsey Nicole Rutherford

A thesis submitted to the faculty of The University of Mississippi in partial fulfillment of the requirements of the Sally McDonnell Barksdale Honors College.

Oxford
May 2021

Approved by

_____
Advisor: Dr. Randy Wadkins

_____
Reader: Dr. Saumen Chakraborty

_____
Reader: Dr. Susan Pedigo

ACKNOWLEDGEMENTS

I would first like to thank my advisor, Dr. Randy Wadkins, for his support and assistance throughout this process. Dr. Wadkins welcomed me into his lab at the end of my sophomore year. These past two years, Dr. Wadkins has answered my many questions and encouraged me throughout my time doing research. I appreciate Dr. Wadkins for giving me the knowledge and guidance to complete this thesis. I would also like to thank my second reader, Dr. Chakraborty, and my third reader, Dr. Pedigo, for taking the time to read my thesis and providing additional guidance.

I would like to thank the Sally McDonnell Barksdale Honors College for the opportunity to write this thesis. The Sally McDonnell Barksdale Honors College has pushed me to grow as a leader and a citizen scholar during my time at Ole Miss. I would like to extend my gratitude to the Chemistry department and faculty for the knowledge and wisdom I have acquired through the countless hours, days, and years I spent in Coulter Hall.

Lastly, I would like to thank my friends and family for their unyielding support and love throughout this entire journey. Without their reassurance and encouragement, I wouldn't be where I am today.

ABSTRACT

LINDSEY NICOLE RUTHERFORD: Chemical and Co-Solute Effects of Polyethylene Glycol on i-Motif Formation
(Under the direction of Randy Wadkins)

DNA typically forms Watson and Crick double helix structures in which adenine, thymine, guanine, and cytosine pair with their complimentary DNA base. However, DNA i-motif structures can form in cytosine rich DNA, typically under slightly acidic conditions (~pH 6). DNA i-motifs are four stranded secondary structures in which cytosine pairs with cytosine to form a quadruplex. The i-motifs are typically formed in acidic conditions because of the protonation in the C•C base pair between one of the three hydrogen bases. Recent studies have suggested i-motifs can also form under neutral conditions, which is more realistic for a cell. It is important to study i-motifs because of their suggested role in gene expression, drug targeting, and drug delivery. The i-motifs can be studied under different physiological conditions and with the addition of crowding agents. The DNA strand DUX4L22 was used to study the formation of i-motifs from a pH range of 5.4 to 7.9 in solutions containing polyethylene glycol (PEG). These conditions mimic those found within a cell. The formation of i-motif structures was observed using circular dichroism spectroscopy. It was found that without the addition of a crowding agent, i-motif formation occurred at only acidic conditions. With the addition of 5% PEG 300, i-motif formation occurred at pH 7.1. With the addition of 20% PEG 300, i-motif formation occurred at pH 7.5. Lastly, with the addition of 30% PEG 300, i-motif formation occurred at pH 7.9. Therefore, it was concluded that

with the addition of crowding agents, DNA i-motif formation can occur in basic physiological conditions for DUX4L22.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF ABBREVIATIONS

Deoxyribonucleic acid – DNA
i-Motif- iM
Polyethylene Glycol 300 – PEG 300
Circular Dichroism – CD
Ultraviolet- Visible – UV-Vis
Tris-EDTA – TE
Nanometer- nm
Microliter - μL
Milliliter- mL
Sodium Cacodylate Buffer - SCB

INTRODUCTION

Deoxyribonucleic acid, or DNA, is the building block of all life. DNA was first discovered in the late 1860s by a chemist named Friedrich Miescher (Pray, 2008). During experiments on leucocytes, Miescher isolated a molecule unlike the proteins he was planning to study. He termed the molecule "nuclein" because of its location in the nucleus (Dahm, 2005). This groundbreaking discovery wasn't seen as important as it should have back in the 19th century and went quiet for nearly 50 years. It was not until the 1950s that more information was determined about the "nuclein", later to be renamed deoxyribonucleic acid, or DNA.

After the discovery that DNA was genetic material in the 1940s by scientists Levene and Chargaff, the structural components were studied by Watson and Crick who pioneered the double helix that most people recognize today. DNA is composed of "two long polynucleotide chains" (Alberts et al., 2002). Within each chain there are nucleotides. Nucleotides are composed of the sugar deoxyribose, a phosphate group, and a base. There are four nucleotide bases- adenine, cytosine, guanine, and thymine. In Watson and Crick DNA, the nucleotide bases pair with one another. Specifically, adenine pairs with thymine and cytosine pairs with guanine by forming hydrogen bonds. These bases bond between the polynucleotide chains to connect the two. Watson and Crick saw that each strand was composed of phosphates and sugars on the outside and bases connected by hydrogen bonds in the inside (Alberts et al., 2002). In addition, the distance

between the two strands is equal throughout the DNA, and DNA winds to form a double helix to base-stack efficiently (Alberts et al., 2002). The DNA is also complementary as one strand ends in a hydroxyl and the opposite ends in a phosphate. This is the typical structure of DNA discovered by Watson and Crick and is shown in **Figure 1**. However, there are other DNA structures that can form such as a DNA i-motif.

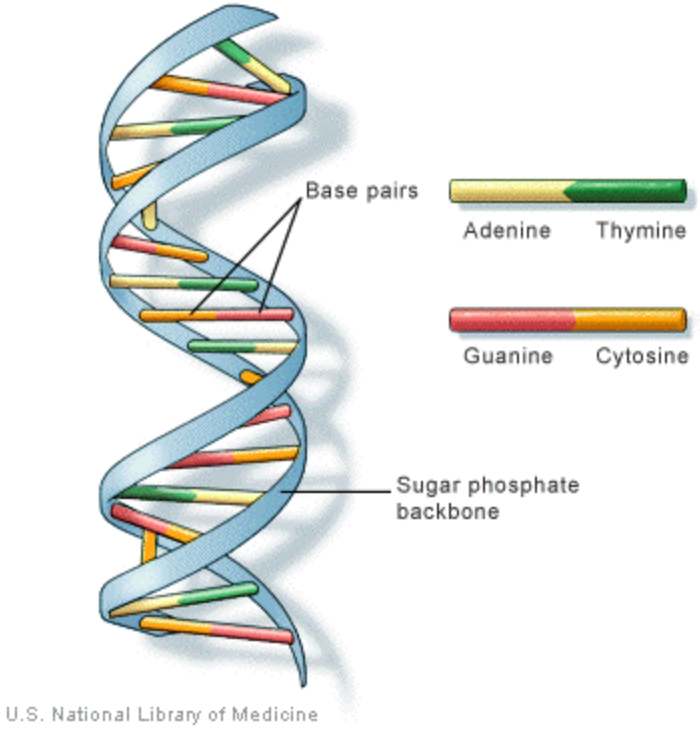


Figure 1: Double-helix DNA structure discovered by Watson and Crick.

from Clark, Jim (2016, May). DNA- Structure.
https://www.chemguide.co.uk/organicprops/aminoacids/dna1.html

DNA i-motifs were discovered in 1993 by scientists Gehring, Leroy, and Gueron (Day et al., 2014). Instead of the typical two stranded Watson and Crick DNA, i-motifs, or iMs, are four-stranded and composed primarily of the base cytosine. The cytosine base pairs are hemi-protonated, meaning one cytosine in the pair has a positive charge (Abou Assi et al., 2018). These hemi-protonated cytosine base pairs hold together the four strands with hydrogen bonds and therefore form a tetramer (Abou Assi et al., 2018).

Furthermore, DNA i-motif tetrads can vary in length, ranging from 3 to 12 base pairs, and can form two different tetramers depending on the direction of the hemi-protonated cytosine pair (Day et al., 2014). In other words, the C-C^+ can end on a 3' or 5' end, which will result in two different tetramers. A typical iM structure is shown in **Figure 2** and can be compared to Watson and Crick DNA displayed in **Figure 1**.

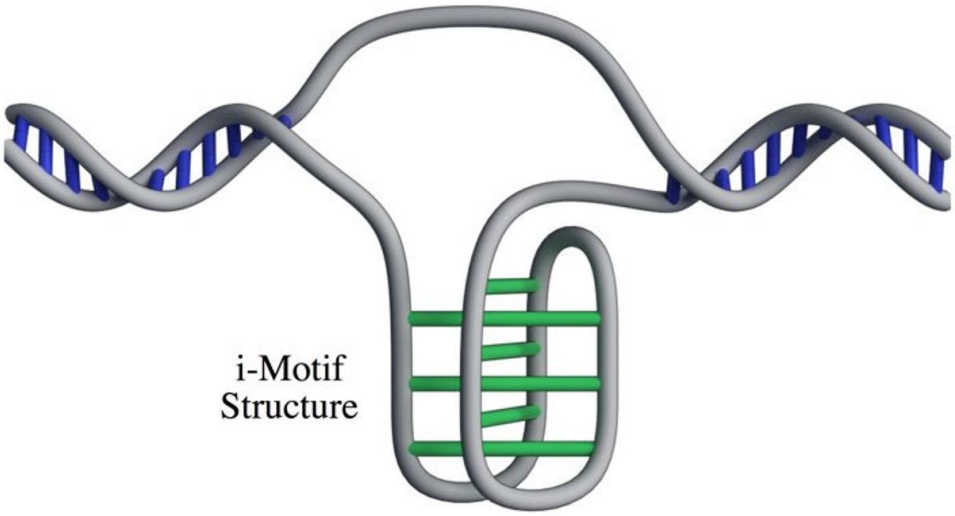


Figure 2: i-Motif Structure found in DNA

From Zeraati, M., Langley, D. B., Schofield, P., Moye, A. L., Rouet, R., Hughes, W. E., et al. (2018). I-motif DNA structures are formed in the nuclei of human cells. Nature Chemistry, 10(6), 631-637. doi:http://dx.doi.org.umiss.idm.oclc.org/10.1038/s41557-018-0046-3

Acidic conditions are preferred for iM stability because of the hemi-protonation states of the cytosine (Abou Assi et al., 2018). However, some studies have begun to show that iMs can also form in more neutral conditions depending on environmental factors. When forming in neutral conditions, a free H+ ion is taken up upon formation of the iM structure (Wright et al., 2017). In addition, iMs are believed to exist *in vivo*, or inside of the cell. If this is the case, it would make more sense for iMs to form at neutral conditions to represent the conditions of *in vivo* (Abou Assi et al., 2018). In a study

performed by Christ et al., it was found that *in vivo* formation of iMs are pH dependent, cell cycle dependent, and are formed in regulatory regions of the human genome (Zeraati et al., 2018). Their study involved using an antibody fragment that was specific to an iM telomere and analyzed the kinetics and affinity of the antibody fragment with the iM. It was found that in addition to iMs forming in telomeric regions, iM structures *in vitro* are found in centromeres and proto-oncogenes which could help explain the biological relevance of iMs (Zeraati et al., 2018). This was found by co-staining cells using the antibody fragment iMab and the telomere repeat-binding factor 2 (TRF2). Microscopy 3D images were observed to detect localization of the staining (Zeraati et al., 2018). The co-localization of iMab and TRF2 suggested that i-motif structures are likely to exist on the foci TRF2 of telomeres (Zeraati et al., 2018). Overall, it was suggested from this study that DNA iMs are found in regulatory regions of the genome. These iMs are believed to play a role in gene regulation (Sedghi et al., 2018). In addition, they can therefore act as a molecular switch within gene expression and drug targeting (Benabou et al.,2014). Their believed function can be useful to serve as a biological target for the development of drugs (Wright et al., 2017). In fact, their function could possibly play a role in tumor formation and could be targeted by antitumor drugs. There is still much that is not known about the role and function of iMs, and that is why the research in this thesis is significant.

The formation of iMs under neutral conditions has been found to be more likely with the addition of molecular crowding agents. This represents cellular conditions that would be realistic for iM formation, since the eukaryotic cell is extremely crowded. High molecular weight polyethylene glycols (PEG) have been used to mimic a crowded

environment and preferentially stabilize iM structures (Abou Assi et al., 2018). In this thesis, Polyethylene Glycol 300 was used as molecular crowding agent to represent *in vivo* conditions.

Polyethylene Glycol 300, or PEG 300, is a macromolecule composed of polymeric ethylene glycol monomers (Moore, 2020). It has many properties similar to sugars. **Figure 3** shows the structure of PEG 300. "PEG and other crowding agents can increase the rate of enzymatic reactions, alter reaction products, protect macromolecules from thermal denaturation, accelerate protein folding, and facilitate nucleic acids renaturation" (Akabayov et al., 2013). The long chains of PEG can act both to exclude water and to act as barriers to diffusion of macromolecules, hence the "crowding" effect. The purpose of adding the PEG 300 was to analyze whether it affected iM formation.

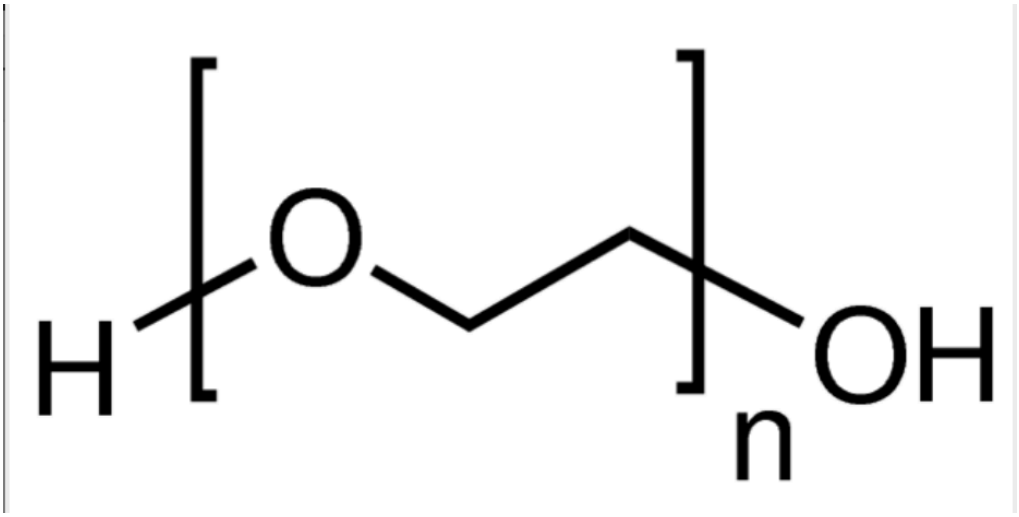


Figure 3: Chemical Formula for Polyethylene Glycol where M_n is 300.

From Millipore Sigma. Polyethylene Glycol 300. https://www.sigmaaldrich.com/catalog/product/mm/807484?lang=en&region=US&cm_sp=Insite-_-caSrpResults_srpRecs_srpModel_peg%20300-_-srpRecs3-1

To study the formation of iMs, circular dichroism, or CD, spectroscopy can be used. Circular dichroism is a type of absorption spectroscopy that measures the difference

in right and left polarized light (Johnson, 1996). This difference can be measured because optically active chiral molecules usually prefer either left or right polarized light (Hurlburt, 2020). To understand the concept of CD spectroscopy, one must first understand the polarization of light. “Electromagnetic radiation is generally described as a transverse wave resulting from an oscillating electric field orthogonal to the propagation direction” (Drake, 1986). If the oscillating electric field vector transcribes a helical path that can be left or right, the light is circularly polarized (Drake, 1986). Circularly polarized light can be either left circularly polarized or right circularly polarized. Left circularly polarized light rotates counter clockwise and right circularly polarized light rotates clockwise (Hurlburt, 2020). This light can be absorbed by molecules based on concentration, temperature, pH and other environmental factors and can then be measured through instrumentation (Hurlburt, 2020). A CD spectrometer passes singular linearly polarized light through a monochromator and then to a photo-elastic modulator which changes the light from linear to circular (Drake 1986). As the incident light switches between left and right circularly polarized light, the absorption changes and molar absorptivity can be measured through a detector (Hurlburt, 2020). **Figure 4** shows a schematic for a CD spectrometer.

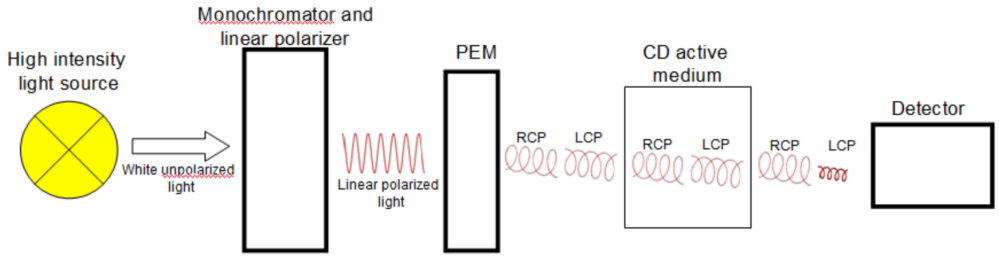


Figure 4: Schematic for Circular Dichroism Spectrometer

From Hurlburt, Nick. (2020, August 15). Circular Dichroism. https://chem.libretexts.org/Bookshelves/Physical_and_Theoretical_Chemistry_Textbook_Maps/Supplemental_Modules_(Physical_and_Theoretical_Chemistry)/Spectroscopy/Electronic_Spectroscopy/Circular_Dichroism.

Circular dichroism is a widely used technique for studying the secondary structures of proteins because it requires low sample concentrations and allows for many variables to be isolated and studied, such a temperature and pH (Correa & Ramos, 2009). Particularly, circular dichroism is a good method for studying and analyzing iM structures because CD scans can also identify structural components of DNA in addition to proteins (Polavarapu & Zhao, 2000). The values displayed in a CD spectrum show the conformational change that occurs in a molecule. In this case, the CD spectrum shows if an iM occurs or does not occur.

In addition to CD spectroscopy, Ultraviolet-Visible, or UV-Vis, spectroscopy was used to determine the concentration of DNA present in a sample. UV-Vis spectroscopy is a quantitative technique that measures how much a chemical absorbs light (Edinburgh Instruments, 2021). This is done by measuring the intensity of light as it passes through a sample compared to the intensity of light without the sample (Edinburgh Instruments 2021). Determining the absorbance can allow one to then determine the concentration of a sample. This is done by using the Beer-Lambert Law. The Beer-Lambert law relates

electromagnetic radiation to the path length and concentration of the absorbing species (Swinehart, 1962). The Beer-Lambert Law is represented in **Equation 1**.

$$A = Ebc$$

A= absorbance, E= molar absorptivity, b= path length, and c= concentration.

Equation 1: The Beer Lambert Law

From Swinehart, D. (1962). The Beer Lambert Law. *Journal of Chemical Education, 39*.

doi:https://pubs.acs.org/doi/abs/10.1021/ed039p333

The DNA strand DUX4L22 was used for the analysis of iM formation. The DUX4L22 sequence has not been extensively studied. The Waller Lab conducted research on DUX4L22 that reported iM formation at pH~7 and concluded that iM formation cannot be excluded in neutral conditions (Wright et al., 2017). The DUX4L22 sequence is 5'- CCC CCG AAA CGC GCC CCC CTC CCC CCT CCC CCC TCT CCC CC – 3'. As shown, DUX4L22 is composed of primarily cytosine. In fact, the G+C content for this strand was 82.9%. Therefore, the DUX4L22 DNA strand is very useful for studying DNA i-Motif formation. The molecular weight is 12,046.8 g/mol. The sample of DNA in this experiment was commercially prepared and received dry as 642.1 nmol of DNA. The extinction coefficient for this strand is 321,700 L/(mole·cm). This information is relevant when applying Beer's Law and when creating the buffer solutions.

METHODS

Materials

DNA strand - DUX4L22:

5'-CCC CCG AAA CGC GCC CCC CTC CCC CCT CCC CCC TCT CCC CC -3'

Tris-EDTA (TE) buffer pH 8

1.0 M HCl

1.0 M NaOH

Cacodylic Acid, various pH solutions

PEG 300

Preparation of DNA

The DNA strand DUX4L22 was obtained from IDT Technologies. 642.1 nmol of DNA was present. TE buffer at pH 8 was added as a storage buffer for the DNA. 642.1 μL was added to the tube to create a 1 mM solution. This solution was then centrifuged and could be stored for further use.

Sodium Cacodylate Buffer Preparation

A 30 mM sodium cacodylate buffer was prepared to manipulate pH conditions of the DNA solutions made. Cacodylic acid has a molecular weight of 214.03 g/mol. Therefore, to create a 30 mM buffer, 6.42 g was dissolved in 1 L of deionized water. The solution

was then mixed using a stir bar until all cacodylic acid was dissolved. This buffer was then divided into ten 15 mL tubes. The pH wasted adjusted by using a pH meter. To make the solution more acidic, 1.0 M HCl was added dropwise. To make the solution more basic, 1.0 M NaOH was added dropwise. Ten solutions were made with pHs 5.4, 5.8, 6.0, 6.2, 6.4, 6.6, 6.8, 7.1, 7.5, and 7.9.

PEG 300 Buffer Preparation

To analyze the effects of a crowding agent, PEG 300 was introduced into the sodium cacodylate buffers. Solutions of 5%, 10%, 20%, 30%, and 40% PEG 300 were made. These were made by using the following calculations. For 40% PEG, 40 g of PEG 300 could be added to 60 g of sodium cacodylate buffer. The 40 g of PEG 300 corresponded to ~33.33 mL and 60 g of sodium cacodylate buffer corresponded to ~54.5 mL. These amounts were then reduced by a factor of 5 to have smaller total volumes. Therefore, 6.66 mL of PEG 300 was added to 10.9 mL of sodium cacodylate buffer for the 40% solutions. For 30% PEG 300, 5 mL of PEG 300 was added to 12.73 mL of sodium cacodylate buffer. For 20% PEG 300, 3.33 mL of PEG 300 was added to 14.55 mL of sodium cacodylate buffer. For 10% PEG 300, 1.67 mL of PEG 300 was added to 16.36 mL of sodium cacodylate buffer. Lastly, for 5% PEG 300, 1.04 mL of PEG 300 was added to 21.59 mL of sodium cacodylate buffer. Ten solutions of PEG 40%, 30%, 20%, 10%, and 5% were made and the pH of each was then adjusted accordingly. For each grouping of PEG, there was a solution of pH 5.4, 5.8, 6.0, 6.2, 6.4, 6.6, 6.8, 7.1, 7.5, and 7.9.

DNA Solutions with Buffers

Before running CD scans, DNA solutions were made consisting of the desired buffer and DNA stock solution. All samples were 5 μM and were made by adding 5 μL of DNA stock to 995 μL of buffer. It is important to add the larger volume first followed by the addition of DNA. The buffer used was either sodium cacodylate buffer alone or with PEG. Prior to adding the PEG buffer, the buffers were filtered through a 0.45 μM syringe filter to ensure no contamination or debris. This solution was then mixed by hand and put into a heating block with temperature 80 °C for 5-10 minutes. This allowed the DNA to spread evenly throughout the solution and removed any mismatches. Once heated, the solution was cooled to room temperature before performing UV-Vis and CD scans.

UV-Vis Scans

UV-Vis scans were performed to analyze the absorbance and concentration of the DNA solutions. The instrument used was an Ocean Optics UV-Vis. DNA solutions were transferred to a cuvette and placed in the spectrometer. The absorbance was then measured at ~260-280 nm and recorded. The Beer-Lambert law could then be applied to the found absorbance to determine the exact concentration of DNA present in the solution. This was performed prior to each CD scan so that CD spectra could be scaled to DNA concentration.

CD Scans

A CD scan was performed to determine if iM formation was present in the DNA solution at designated pHs with and without the addition of crowding agents. An Olis DSM-20

CD instrument was used to conduct CD scans. After turning on the instrument, the Olis application was opened. The temperature was set at 25°C. CD scans were run from a wavelength of 350 nm to 220 nm with 1 nm increments. A digital filter was used for smooth graphs. First, a scan was collected of just buffer. The buffer was added into a cuvette and placed in the spectrometer. Data was collected and the buffer solution was assigned as the baseline. The cuvette was emptied and replaced with the DNA buffer solution. Then, the cuvette was placed in the spectrometer and data was collected. Once the scan was complete, the data was exported to excel and graphed through LoggerPro. After all scans were completed, the cuvette was cleaned with deionized water and methanol.

RESULTS AND DISCUSSION

To determine if i-motif formation was present, circular dichroism graphs were analyzed. The iM graphs are distinctive when compared to DNA that does not have iM formation. Unordered DNA that is C-rich typically has a maximum peak ~275 nm on the spectrum (Školáková et al., 2019). As pH decreases below the neutral level (pH<7), the maximum peak shifts to higher wavelength, known as a red shift (Školáková et al., 2019). Simultaneously, the region ~260 nm deepens which creates a negative band (Školáková et al., 2019). The CD spectrum then has a maximum positive value ~287 nm and a minimum negative value ~265 nm which shows that an iM is present (Školáková et al., 2019). Therefore, these characteristics were examined in each spectrum. When present, iM formation was said to have occurred.

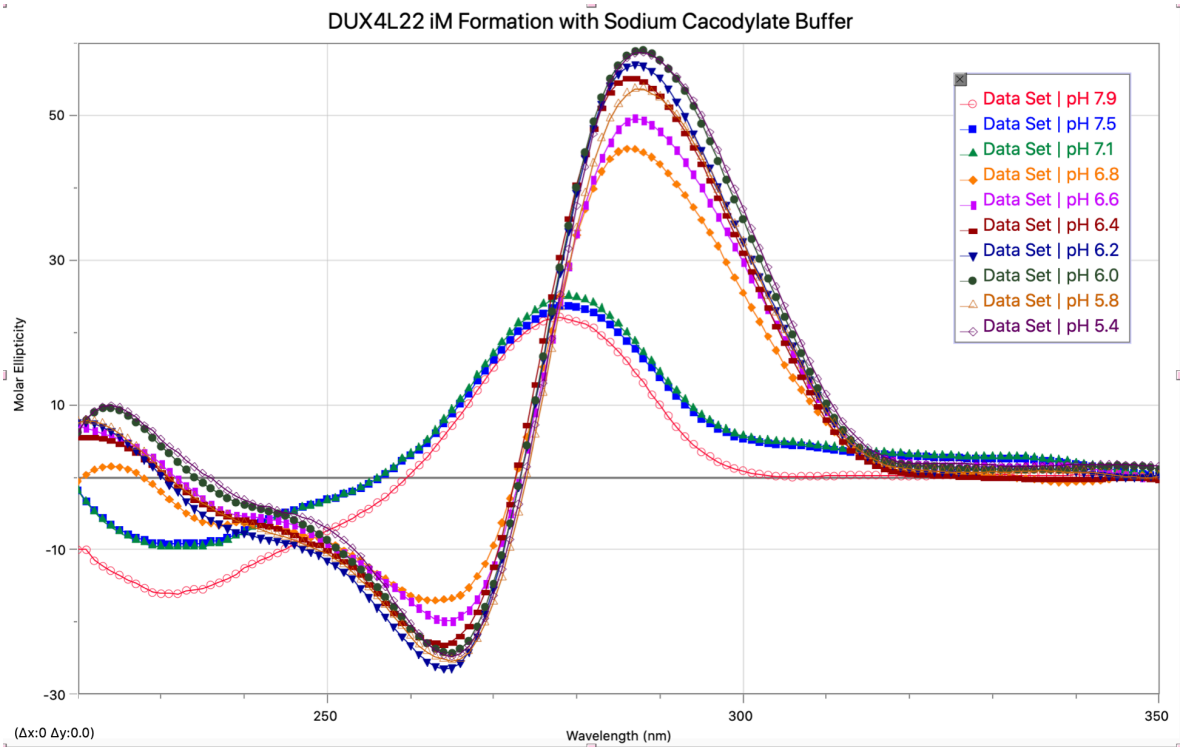
DNA iM formation was first analyzed with just the addition of sodium cacodylate buffer. Since no crowding agents were present, it was predicted that iM formation would occur at acidic conditions (pH<7.0). For each CD scan, the concentration of DNA was accounted for by determining the molar ellipticity. **Table 1** contains the measured absorbances for all DNA solutions used during experimentation. **Table 2** contains the calculated concentrations for all DNA solutions. In addition, curves were adjusted for negative starting values. **Graph 1** shows the data collected for pH range of 5.4-7.5 for the DNA strand DUX4L22 in the absence of PEG.

Absorbances						
pH	DNA w/SCB	5% PEG	10% PEG	20% PEG	30% PEG	40% PEG
7.9	1.232	1.13	1.694	1.472	1.282	1.419
7.5	1.267	1.252	1.307	1.67	1.322	1.439
7.1	1.205	1.158	1.269	1.455	1.337	1.485
6.8	0.852	1.244	1.519	1.392	1.464	1.467
6.6	1.249	1.297	1.439	1.254	1.405	1.522
6.4	1.279	1.485	1.549	1.933	1.443	1.569
6.2	1.235	1.475	1.583	1.876	1.497	1.802
6.0	1.209	1.517	1.613	1.423	1.534	2.147
5.8	1.315	1.862	1.774	1.465	1.481	1.834
5.4	1.298	1.472	1.686	1.47	1.621	1.744

Table 1: Absorbances measured by UV-Vis Spectroscopy.

Concentrations (mM)						
pH	DNA w/ SCB	5% PEG	10% PEG	20% PEG	30% PEG	40% PEG
7.9	0.0038	0.0035	0.00527	0.00458	0.00399	0.00441
7.5	0.0039	0.0039	0.00406	0.00519	0.00412	0.00447
7.1	0.0037	0.0036	0.00394	0.00452	0.00416	0.00462
6.8	0.00265	0.00387	0.00472	0.00432	0.00455	0.00457
6.6	0.00388	0.004	0.00447	0.00389	0.00437	0.00473
6.4	0.00398	0.0046	0.00482	0.00601	0.00449	0.00488
6.2	0.00384	0.00459	0.00492	0.00583	0.00465	0.0056
6.0	0.00376	0.00472	0.00501	0.00442	0.00477	0.00667
5.8	0.00409	0.00579	0.00551	0.00455	0.0046	0.0057
5.4	0.00403	0.00458	0.00524	0.00457	0.00504	0.00542

Table 2: Calculated concentrations by using the Beer-Lambert Law.

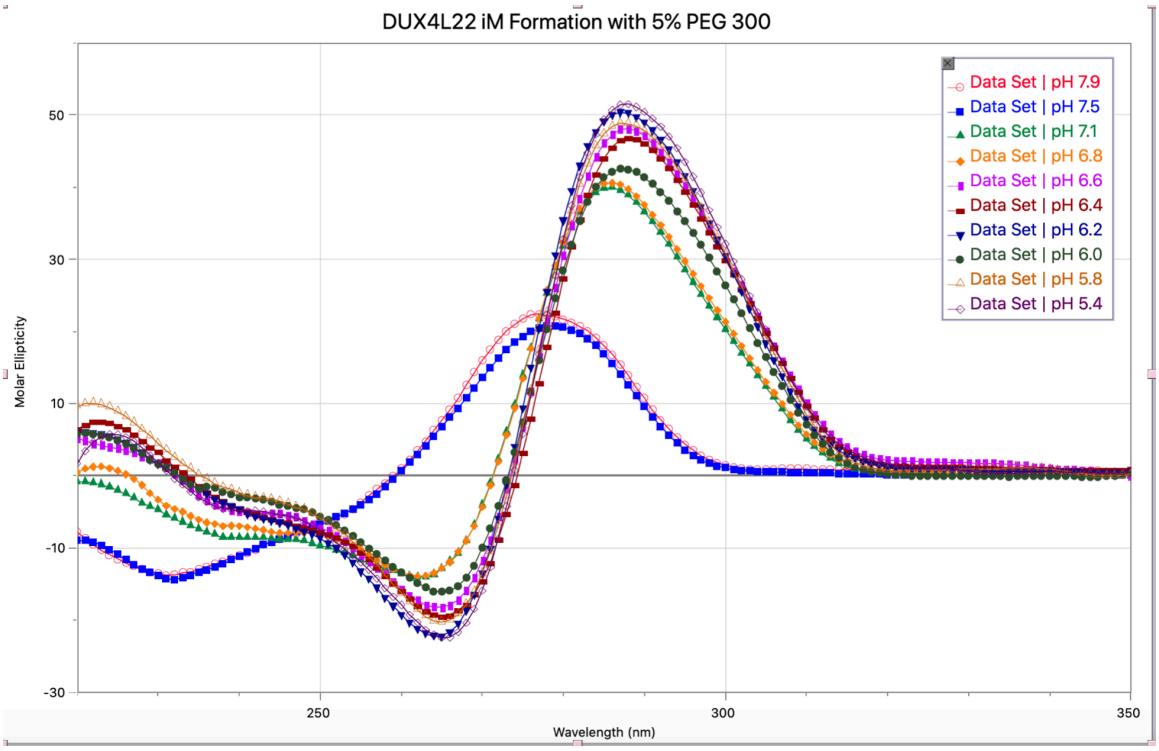


Graph 1: CD spectrum for DNA DUX4L22 with Sodium Cacodylate Buffer pH range 5.4-7.9.

As shown in **Graph 1**, the curves with a pH of 6.8, 6.6, 6.4, 6.2, 6.0, 5.8, and 5.4 have maximum peaks around 287 nm and have negative minimum peaks around 263 nm. These are consistent with iM formation. In contrast, at pH 7.9, 7.5, and 7.1 maximum values occurred around 278 nm. In addition, at pH 7.9, 7.5, and 7.1 the peaks are significantly less intense and have a different shape than the acidic pH values. Therefore, with no addition of crowding agents iM formation occurred only at acidic pH values. Specifically, iM formation occurred at pH 6.8, 6.6, 6.4, 6.2, 6.0, 5.8, and 5.4.

DUX4L22 was then analyzed under conditions with the addition of crowding agents. First, 5% PEG 300 was introduced into the sodium cacodylate buffers. The

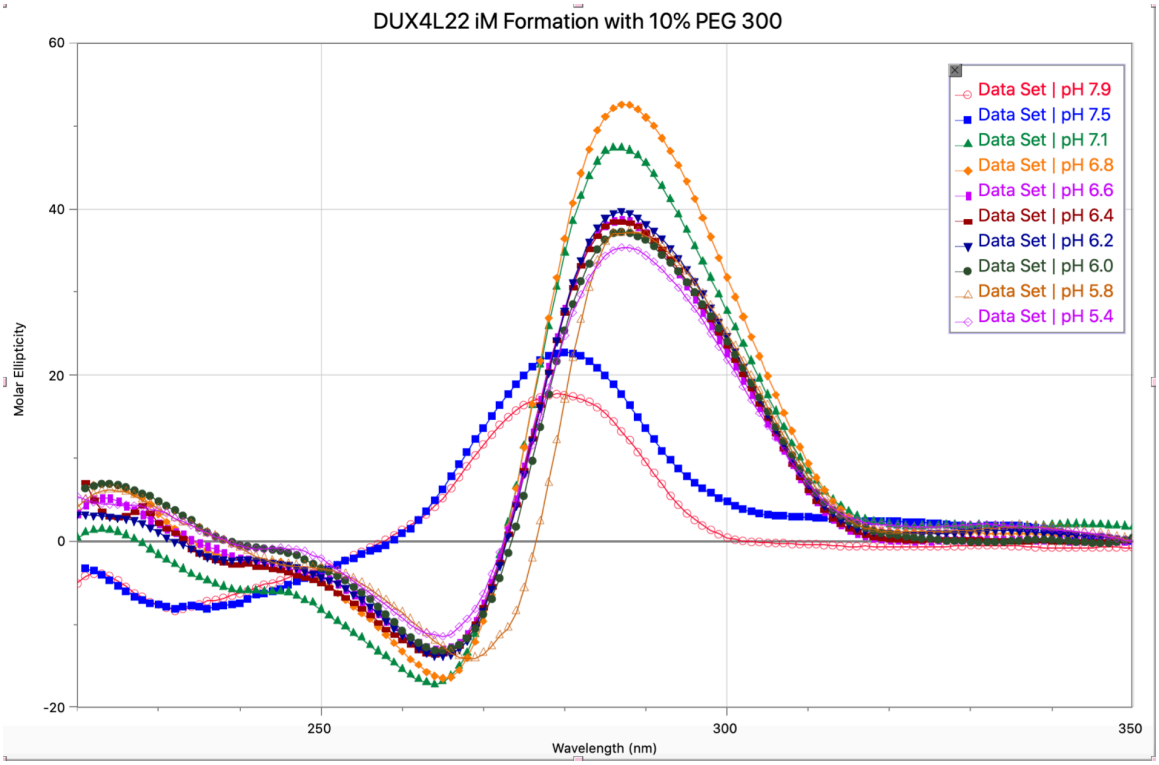
buffers remained clear but became thicker with the addition of PEG 300. **Graph 2** shows the data collected for 5% PEG 300.



Graph 2: CD spectrum for DNA DUX4L22 with the addition of 5% PEG 300 to Sodium Cacodylate Buffers with a pH range 5.4-7.9.

As shown in **Graph 2**, the curves with a pH ≤ 7.1 demonstrate i-motif formation because they contain maximum peaks around 287 nm and minimum peaks around 265 nm. This is significantly important because when compared to **Graph 1,** an i-motif now forms at pH 7.1. Previously, an i-motif did not form in the absence of PEG. This is likely due to the addition of 5% PEG 300, a crowding agent, because no other variables were changed during experimentation. DNA at pH 7.5 and 7.9 still did not form i-motifs as shown by their curves in **Graph 2**. At pH 7.5 and 7.9 maximum values occurred around 275 nm, and therefore, signify no i-motif formation. For further analysis, the percent

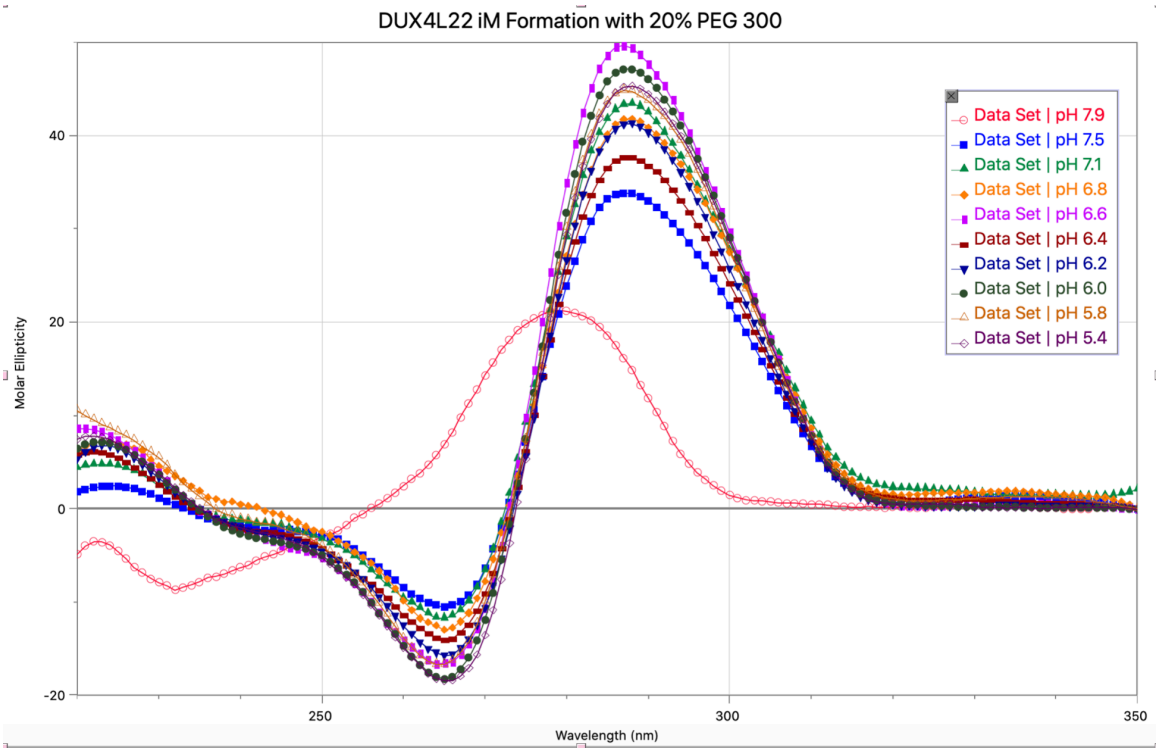
weight of PEG 300 was increased to 10% and analyzed. **Graph 3** shows the CD spectrum collected for 10% PEG 300.



Graph 3: CD spectrum for DNA DUX4L22 with the addition of 10% PEG 300 to Sodium Cacodylate Buffers with a pH range 5.4-7.9.

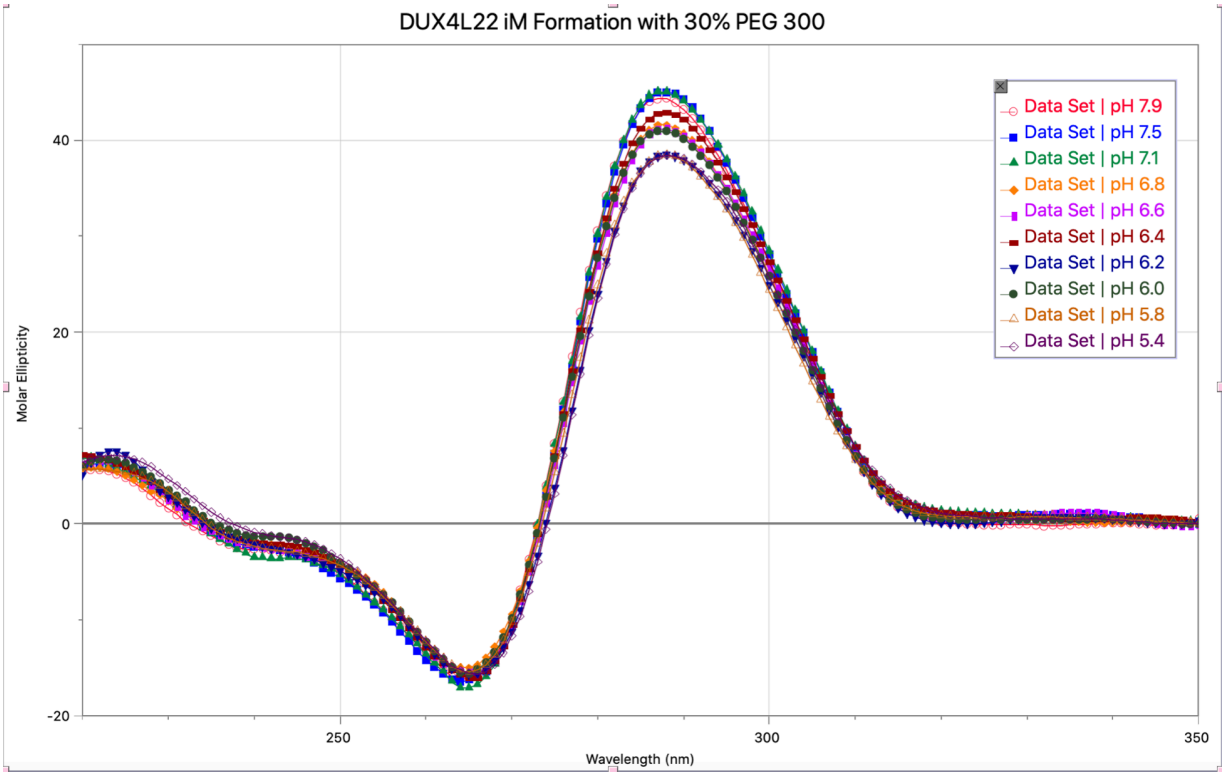
Similarly to **Graph 2**, **Graph 3** shows i-motif formation at pH ≤ 7.1 because the curves have characteristics of i-motif formation. This data is still significant because it provides support that the addition of crowding agents affects the stability and formation of i-motifs. DNA i-motif formation was still not present at pH 7.5 and 7.9 which is understandable since i-motif formations prefer acidic condition. This graph was corrected for concentration yet the graphs still do not overlap. This could be due to systematic or random error when running the CD scans. There could also be energy shifts due to the addition of PEG 300 since peaks occur around different wavelengths. These shifts could

be further studied, and with multiple trials, precision could be determined. After the addition of 10% PEG 300, scans were collected with 20% PEG 300 and are shown in **Graph 4**.

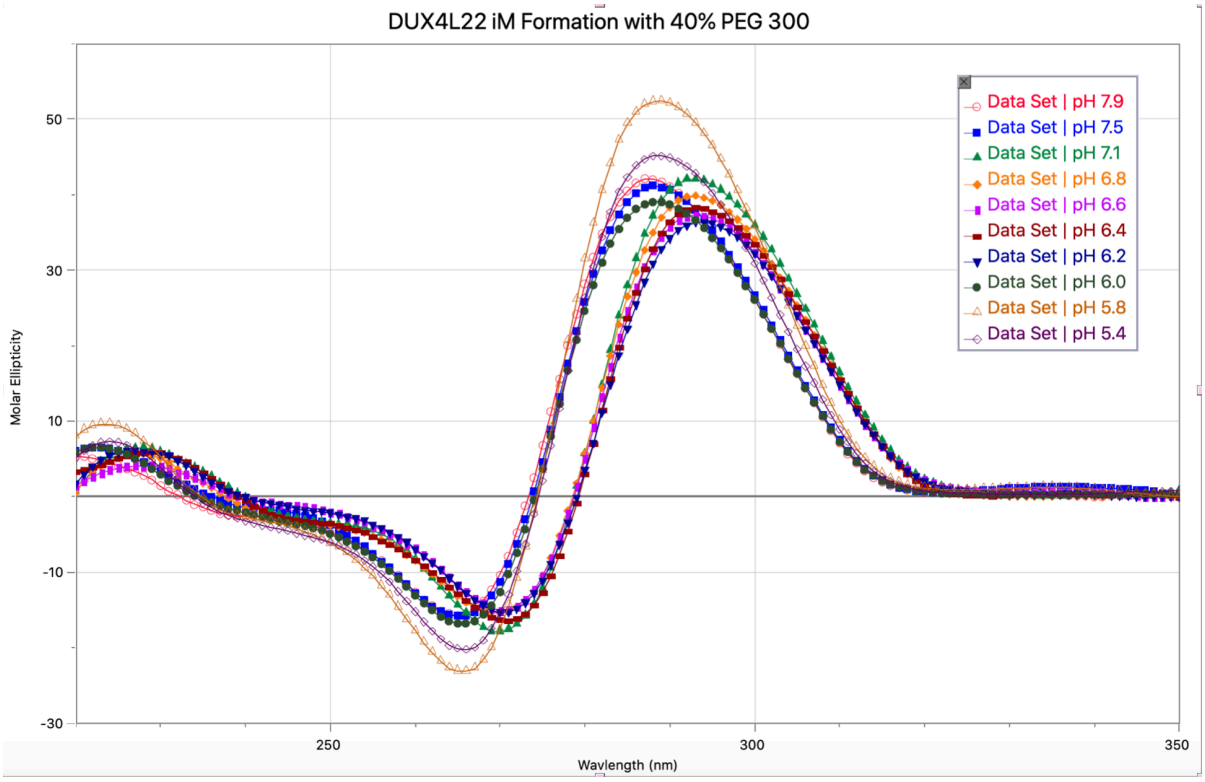


Graph 4: CD Spectrum for DUX4L22 with the addition of 20% PEG 300 to Sodium Cacodylate Buffers with a pH range 5.4-7.9.

As shown in **Graph 4,** i-motif formation is present at pH 7.5 as more crowding agent is added to the solution. This is represented by characteristic i-motif formation maxima and minima. i-Motif formation at pH 7.5 was not previously present due to small amounts of crowding agent. As more crowing agent becomes present, DNA begins to form i-motifs at more basic pH. The i-motif formation remains present at acidic pHs as shown in **Graph 4** but is still absent at pH 7.9. It was predicted that as more crowding agent is added, i-motif formation will likely occur at pH 7.9. The 30% PEG 300 and 40% PEG 300 scans are shown in **Graph 5** and **Graph 6.**

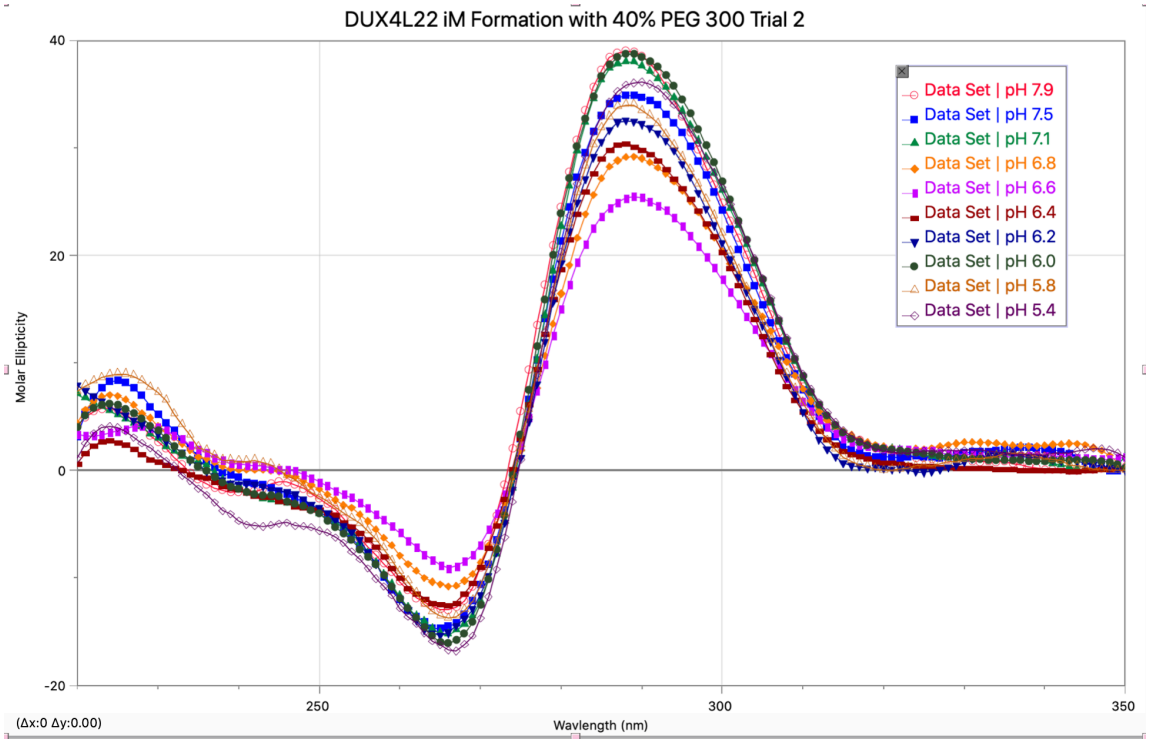


Graph 5: CD spectrum for DNA DUX4L22 with the addition of 30% PEG 300 to Sodium Cacodylate Buffers with a pH range 5.4-7.9.

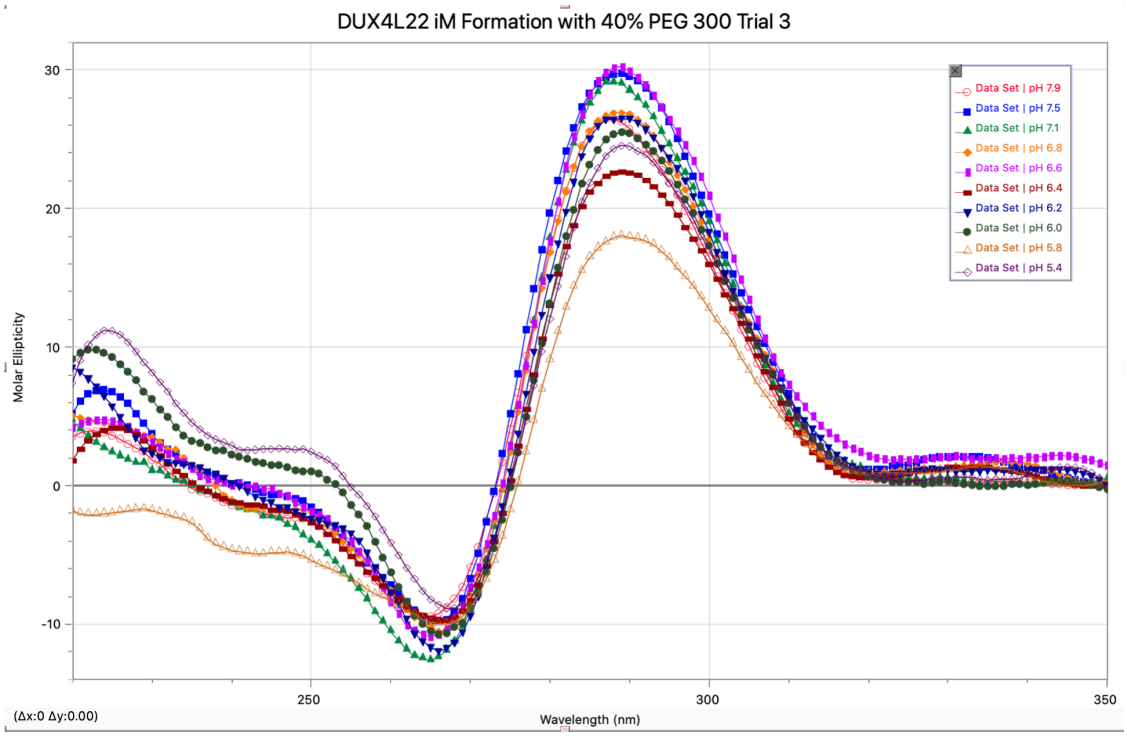


Graph 6: CD spectrum for DNA DUX4L22 with the addition of 40% PEG 300 to Sodium Cacodylate Buffers with a pH range 5.4-7.9.

Both **Graph 5** and **Graph 6**, show i-motif formation at all pH values measured. With a higher percent by weight crowding agent, i-motif formation began to occur at basic pHs as represented by formation in pH values greater than 7.0. As seen in **Graph 6** for 40% PEG 300, a shift in energy is present between varying pH values. This cause could not be explained because DUX4L22 is not well studied. Therefore, two more trials were conducted on 40% PEG 300 to determine if this shift was human error or if a shift in energy is present. In addition, it was important to carry out more trials to ensure that formation was occurring at pH 7.9. **Graph 7** and **Graph 8** show the CD spectrum for trial 2 and trial 3 of 40% PEG 300 scans. **Table 3** contains the absorbances and concentrations for the DNA strands measured.



Graph 7: Second trial for DNA DUX4L22 with the addition of 40% PEG 300 to Sodium Cacodylate Buffers with a pH range 5.4-7.9.



Graph 8: Third trial for DNA DUX4L22 with the addition of 40% PEG 300 to Sodium Cacodylate Buffers with a pH range 5.4-7.9.

Trial 2		Trial 3	
Concentration (mM)	Absorbance	Concentration (mM)	Absorbance
0.00177	0.00244	0.572	0.784
0.00197	0.00206	0.635	0.662
0.00194	0.00223	0.624	0.718
0.00256	0.00235	0.823	0.755
0.00318	0.00213	1.022	0.686
0.00255	0.00276	0.821	0.888
0.00201	0.00290	0.647	0.671
0.00354	0.00247	1.139	0.796
0.00193	0.00343	0.622	1.104
0.00189	0.00252	0.608	0.811

Table 3: Concentration and Absorbances for Trial 2 and 3 for DUX4L22.

Graph 7 and **Graph 8** show that i-motif formation is still present at pH 7.9. However, when compared to **Graph 6**, a shift in energy is not present between varying pH values. Therefore, it can be concluded that the shift in energy in **Graph 6** was due to error. In contrast, the consistency of i-motif formation at pH 7.9 between trials 1-3 shows that the data is accurate.

This data is significant and can allow researchers to better understand the environment in which i-motif formation occurs. Since crowding agents mimic a cellular environment, it can be predicted that i-motifs may form at neutral and basic conditions inside the cell. The graphs and data presented also show a correlation between the addition of crowding agents and i-motif formation. The more crowding agents that were added resulted in i-motif formation at more basic pH values. For further analysis of 40% PEG 300, more trials can be conducted in the future and compared to other DNA strands that are likely to form i-motifs. Furthermore, i-motif formation for DUX4L22 could be studied with the addition of other crowding agents such as Ficoll and Dextran.

CONCLUSION

DNA i-motif structures are suggested to play a role in biological processes such as gene regulation in addition to their potential use in nanotechnology. It is important to study their formation to better understand and analyze their function. There has been much research on DNA i-motifs in acidic conditions but little is known about their formation in basic conditions. Because of the protonation state of cytosine, acidic conditions are favored (Abdelhamid & Waller, 2020). However, with the addition of molecular crowding agents, i-motif formation in neutral and basic conditions becomes observable. The purpose of this study was to analyze DNA i-motif formation under different conditions that are similar to those found in cells for the strand DUX4L22. i-Motif formation was analyzed with the molecular crowding agent PEG 300. CD scans were studied and graphs that contained i-motif characteristics were identified. It was found that when PEG 300 was not present, i-motifs only formed in acidic conditions. When 5% PEG 300 was introduced into the solution, an i-motif formed at pH 7.1. Furthermore, at 20% PEG 300, i-motif formation occurred at pH 7.5. Lastly, with 30% and 40% PEG 300, i-motif formation occurred at pH 7.9. This signifies that with molecular crowding agents, DNA i-motif formation is possible in neutral and basic conditions. Although the stability of these i-motifs is not known, future research could be conducted to further study their properties. In addition, future research could study other molecular crowding

agents to ensure accuracy and precision. While conducting this research, energy shifts in the CD spectra were seen that could have been a result of the addition of PEG. Since multiple trials were not run for each addition of PEG 300 scans, these shifts could not be verified or studied further. Overall, the goal of this research was achieved and the data collected can be beneficial and useful to future researchers. With expanding and new research on i-motifs, concepts such as drug targeting, drug delivery, and new disease treatments can be developed that include i-motifs.

REFERENCES

Abdelhamid, M. & Waller, Z. (2020). Tricky Topology: Persistence of Folded Human Telomeric i-Motif DNA at Ambient Temperature and Neutral pH. *Front. Chem. 8*:40. doi: 10.3389/fchem.2020.00040

Abou Assi, H., Garavís, M., González, C., & Damha, M. J. (2018). i-Motif DNA: structural features and significance to cell biology. Nucleic acids research, 46(16), 8038–8056. https://doi.org/10.1093/nar/

Akabayov, B., Akabayov, S. R., Lee, S., Wagner, G., & Richardson, C. C. (2013). Impact of macromolecular crowding on DNA replication. *Nature Communications 4, 1615*. Retrieved February 20, 2021, from https://www.nature.com/articles/ncomms2620.

Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., Walter, P. (2002). Molecular Biology of the Cell. 4th edition. New York: Garland Science. The Structure and Function of DNA. Retrieved February 20, 2021 from https://www.ncbi.nlm.nih.gov/books/NBK26821/

Benabou, S., Avino, A., Eritja, R., Gonzalez, C., & Gargallo, R. (2014). Fundamental Aspects of the nucleic acid i-motif structures. *RSC Advances, 4*, 26956-26980. Retrieved from https://www.semanticscholar.org/paper/Fundamental-aspects-of-

the-nucleic-acid-i-motif-Benabou-Avi%C3%B1%C3%B3/f75c589b006f7f7fd88f075064ff5d3e3d91870c.

Correa, D., & Ramos, C. H. (2009). The use of circular dichroism spectroscopy to study protein folding, form, and function. *African Journal of Biochemistry Research, 3*(5), 164-173. Retrieved March 25, 2021, from https://www.researchgate.net/publication/228351774_The_use_of_circular_dichroism_spectroscopy_to_study_protein_folding_form_and_function.

Dahm, R. (2005). Friedrich Miescher and The Discovery of DNA. *Developmental Biology, 278*(2), 274-288. Retrieved February 5, 2021, from https://www.sciencedirect.com/science/article/pii/S0012160604008231

Day, H. A., Pavlou, P., & Waller, Z. A. (2014). I-Motif DNA: Structure, stability and targeting with ligands. *Bioorganic & Medicinal Chemistry, 22*(16), 4407-4418. Retrieved February 05, 2021, from https://www.sciencedirect.com/science/article/abs/pii/S0968089614004064.

Drake, A.F. (1986). Polarisation modulation-the measurement of linear and circular dichroism. J. Phys. E: Scientific Instruments 19(170). Retrieved February 5, 2021 from https://iopscience.iop.org/article/10.1088/0022-3735/19/3/002/meta

Edinburgh Instruments. (2021). UV-Vis Spectroscopy. Retrieved from https://www.edinst.com/us/techniques/uv-vis-spectroscopy/.

Hurlburt, N. (2020). Circular Dichroism. Retrieved February 20, 2021 from

https://chem.libretexts.org/@go/page/1761.

Johnson W.C. (1996) Circular Dichroism Instrumentation. In: Fasman G.D. (eds) Circular Dichroism and the Conformational Analysis of Biomolecules. Springer, Boston, MA. https://doi.org/10.1007/978-1-4757-2508-7_18

Moore, S. (2020). What is Polyethylene Glycol (PEG)? Retrieved February 20, 2021 from https://www.news-medical.net/life-sciences/What-is-Polyethylene-Glycol-(PEG).aspx.

Polavarapu, P. L., & Zhao, C. (2000). Vibrational circular dichroism: a new spectroscopic tool for biomolecular structural determination. *Fresenius' journal of analytical chemistry*, *366*(6-7), 727–734. https://doi.org/10.1007/s002160051566

Pray, L. A., Dr. (2008). Discovery of DNA Structure and Function: Watson and Crick. Retrieved February 05, 2021, from https://www.nature.com/scitable/topicpage/discovery-of-dna-structure-and-function-watson-397/#:~:text=What%20did%20the%20duo%20actually%20discover%3F&text=Many%20people%20believe%20that%20American,by%20Swiss%20chemist%20Friedrich%20Miescher

Sedghi Masoud, S., Nagasawa, K. (2018). i-Motif-Binding Ligands and Their Effects on the Structure and Biological Functions of i-Motif. *Chemical & pharmaceutical bulletin, 66*(12), 1091–1103. https://doi.org/10.1248/cpb.c18-00720

Školáková, P., Renčiuk, D., Palacký, J., Krafčík, D., Dvořáková, Z., Kejnovská, I., Bednářová, K., & Vorlíčková, M. (2019). Systematic investigation of sequence

requirements for DNA i-motif formation. *Nucleic acids research*, *47*(5), 2177–2189. https://doi.org/10.1093/nar/gkz046

Swinehart, D. (1962). The Beer Lambert Law. *Journal of Chemical Education, 39.* doi:https://pubs.acs.org/doi/abs/10.1021/ed039p333

Wright, E. P., Huppert, J. L., & Waller, Z. (2017). Identification of multiple genomic DNA sequences which form i-motif structures at neutral pH. *Nucleic acids research*, *45*(6), 2951–2959. https://doi.org/10.1093/nar/gkx090

Zeraati, M., Langley, D. B., Schofield, P., Moye, A. L., Rouet, R., Hughes, W. E., Bryan, T. M., Dinger, M. E., & Christ, D. (2018). I-motif DNA structures are formed in the nuclei of human cells. *Nature chemistry*, *10*(6), 631–637. https://doi.org/10.1038/s41557-018-0046-3