Investigating Stability of RNA i-motif

Zara Campos Carrillo De Preciado

A thesis submitted for the degree of Master of Science by Research

University of East Anglia

School of Pharmacy

June 2018

Declaration

This thesis is submitted to the University of East Anglia for the Degree of Master of Science and has not been previously submitted at this or any university assessment or for any other degree. Except where stated, and reference and acknowledgment is given, this work is original and has been carried out by the author alone.

Zara Campos Carrillo de Preciado

To you, our dear Auntie Magda with love,

(1926-2013)

Acknowledgements

I would like to thank my supervisor Dr. Zoe Waller for her support and guidance. Also, to all the members of Waller's group for their help. A very special thanks to Mahmoud for sharing his knowledge with me always. And to Dr. Elise Wright who helped me in every step of my project, not only sharing her knowledge but also for being such a lovely and amazing woman, always there ready to answer any questions without hesitation. Her charismatic personality, compassionate attitude and her spunk have been missing in our group. Also, thanks to my second supervisor Dr. Yilliang Ding for everything. So many thanks to Dr. Myles Cheesman of the Henry Wellcome Laboratories for Biological Chemistry, UEA, for the use of the CD spectrometer. Thanks to Dr. Clevor Trevor Simmons and Dr. Sue Jickells for giving me feedback on my paper. Thanks to everyone who helped me to better all my graphics (Oven, Guy) and all my lads, the finest, The Chemie boys (Leggo-my-eggo, Master G, Rossy Pooh, Wooly bag, Big D and the non-binary species) who made my stay in the hollie-jollie England so enjoyable! I would probably be back at home already without your support including some shenanigans. So many thanks to my family in my country of origin, Peru. I am deeply sorry for neglecting you guys especially to my brother Francisco who was in charge of everything in my absence. I love you with all my heart! Hopefully, we will soon be together again. Dear Mamá, this is a small victory! I hope you are proud of me up there! So many thanks to Peter, my husband, lover and partner in crime for his unconditional love, and for reminding me every single day to follow my dreams no matter what!

Abstract

Cytosine-rich sequences can form 4-stranded structures called i-motifs. i-Motifs are formed by two duplexes intercalated in an antiparallel direction and held together through the hemi- protonated cytosine-cytosine base pairs, and they are stabilized by acidic conditions, which promote C-C base pairing via hemi-protonation of the N3. i-Motif structure can form in both molecules of DNA and RNA, but RNA i-motifs are less stable than their DNA counterparts. This dissertation describes the study of the stability of RNA i-motifs including characterizing 10 cytosine-rich sequences with variable tract length in $[C_nU_3]_3U_n$ at two different pH: 5.5 and pH 7.4. A plant sequence AT5G08230.1 RNA identified in *Arabidopsis thaliana* was characterized at pH 5.8 by mimicking crowding with PEG 8,000, and drought conditions using KCl. From our knowledge, there are no examples of functional i-motif structures in plants, but they are of particular interest with respect to potential i-motif formation knowing that plants often thrive in acidic conditions where i-motifs are stable. By using some biophysical techniques like Ultraviolet (UV), Circular Dichroism (CD), and thermal difference spectra (TDS), we found all our cytosine-rich sequences were shown to be stable at pH 5.5, except C_1U_3 and sequences with 5 or more cytosines per tract length fold into stable i-motif structure. Also, we have found the stability of the folded structures increases as the number of cytosines per tract**.** UV spectroscopy has showed an increase in the melting and annealing temperature for all C-rich sequences when the number of cytosines per tract increase**.** On the other hand, at pH 7.4 RNA i-motif structure was not observed. We have observed that AT5G08230.1 RNA can form i-motif at pH 5.8 where plants normally thrive, and have confirmed the RNA i-motif can be stabilized at 4°C using PEG 8,000 (40%) and KCl (100 mM and 800 mM).

Table of Contents

List of Figures

Chapter 1: Introduction.

Chapter 2: Studying the effect of C-tract length on stability of RNA i-motifs using the sequence $[C_nU_3]_3U_n$ or C_nU_3.

List of Tables

Table 11. UV-melting and annealing profile of AT5G in KCl (10 mM), buffer (50 mM pH 5.8) comparing the T_m, T_a average and standard deviation same RNA conditions for No-PEG, PEG 8,000(20%) and PEG 8,000(40%).

List of Abbreviations

A	adenine
Å	Ångstrom
aa	amino acids
C	cytosine
T_a	annealing temperature
C-kit	or tyrosine-protein kinase is an oncogene promoter that is encode in humans by the KIT gene.
C-Myc	oncogene promoter that code for transcription factors from the Myc family. C-Myc was the first gene discovered in this family.
CD	circular dichroism.
*Δ*Abs	change in absorbance
*Δ*pH	change in pH
*Δ*T°	change in temperature
°C	degrees Celsius
DNA	deoxyribonucleic acid
FLC	Flowering Locus Central repressor of flowering time.
FRET	Fluorescence resonance energy transfer
Gly	glycine
G	guanine
hTeloC	human telomeric region

HUA2	genes are in charge of regulate the function and expression of some essential genes in plants.
HULK1	genes are in charge of regulate the function and expression of some essential genes in plants.
T_m	melting temperature
Met	methionine
PAGE	polyacrylamide gel electrophoreses
PEG	polyethylene glycol
pH	potential of hydrogen
KCl	potassium chloride
PDB	protein data bank
RNA	ribonucleic Acid
mRNA	messenger RNA
tRNA	transfer RNA
rRNA	ribosomal RNA
pH_T	transitional pH
SL5	This group of genes regulate the function and expression of some essential genes in plants.
TDS	Thermal Difference Spectrum
T	thymine
UV	Ultraviolet.
U	uracil
XRD	X-ray diffraction

XRC	X-ray crystallography

CHAPTER 1: Introduction

1.0 Introduction

The existence of diverse secondary structures have been identified like duplexes including B form, A and Z form, and triplexes and quadruplex (1). G-quadruplex are the most studied of these, and they are 4- stranded structures formed from guanine-rich DNA sequences forming a guanine tetrad through Hoogsteen hydrogen bonding and their presence *in vitro* has been demonstrated for some decades (2, 3), the presence of g-quadruplexes in the human genome have been demonstrated (4, 5), have been found in repetitive DNA regions like telomeres (6, 7), and in the promoter region of eukaryotic genes (8, 9). The i-motif structure is another four-stranded structure which is formed from cytosine-rich sequences (10), and the stability depends on important factors like pH (10, 11), sequence (11, 12), cytosine tract length (11, 13), loop length (14, 15) and temperature (11, 16). Previous studies about DNA i-motifs had assumed that they were only able to form under slightly acidic conditions (10, 17) but recent studies have shown that DNA i-motifs are able to form at neutral pH (13) and slightly alkaline conditions pH 7.4 at 4°C (16). i-Motif have been observed in the promoter regions of a considerable number of oncogenes (17-19), its role on regulation of transcription and modulation of gene expression have been observed (20). It is important to mention recent studies about the influence of i-motif on DNA replication suggesting i-motif obstructs this process and repair causing as a consequence genomic instability (21). Moreover, the evidence of i-motif structure *in vivo* was just demonstrated in the nuclei of human cells confirming their pH dependency, and evidencing its formation in the regulatory regions of the human genome (22). The RNA i-motif structure has not been widely studied, and the literature around this theme is very limited. RNA i-motifs are known to be less stable than their DNA counterparts (23, 24). Moreover, to the best of our knowledge, until now, there are no examples of functional i-motif structures in plants but they are of particular interest with respect to potential i-motif knowing that plants often thrive in acidic conditions, 76% of arable soils are acidic (25). According to Bradshaw *et. al.* plants should tolerate wider changes in the habitat conditions than animals because they cannot to avoid the cause of stress (26), and plants have to develop mechanism due to changes in environmental conditions like salinity, drought, temperature as a result plants show wider range of pH, temperature and salinity tolerance (27).

All of these factors can affect i-motif structure (28). Martiniere *et. al.* have observed the intracellular pH in the endomembrane system lumen in plants have shown pH gradients from acidification at pH 6.1 to alkaline pH 6.6 and 7.1 in both tobacco (*Nicotiana tabacum*) and *Arabidopsis thaliana* (29). Drought and salinity are some of the main causes of reduction of plant growth around the world. Plants cells under conditions of drought or in high salinity soils avoid cellular dehydration by increasing their ionic concentrations (particularly potassium cations) dramatically (30, 31). There are few studies about the role of external pH on plants gene expression, one of them made in *Arabidopsis thaliana* have confirmed that pH is one of the most important factor regulating gene expression in plants (32). Understanding plant physiology will help protect crops from drought, cell damage and stress. We have postulated a potential role in RNA i-motifs structures which could have switch-like functions activated by conditions to help plants to cope with stress. Bearing this in mind, we propose a project that would aim to investigate the RNA i-motif structure and function in ten cytosine-rich sequences and in a plant sequence to characterize and to determine i-motif stability, and the response to different environmental conditions such as temperature, pH and molecular crowders.

1.1 DNA

The three-dimensional structure of DNA proposed by Watson and Crick is formed by two twisted strands of nucleotides bases joined through hydrogen bonds forming a double helix shape (33) (Figure 1.1). The structure of the molecule of DNA has primary, secondary and tertiary structure. While the primary structure of the molecule of DNA corresponds to arrangement of the nucleotides joined by phosphodiester bonds (34, 35), the secondary structure defines the spatial configuration of double helix of Watson and Crick (33), and the tertiary structure corresponds to the arrangement of the DNA structure in the chromosomes in the nuclei (35).The double helical spatial arrangement of DNA emerged from the properties of its chain of polynucleotides where the bases pair with their respective complementary counterparts in an antiparallel fashion. The sugar pentose in DNA is the deoxyribose sugar attached to a phosphate group that based on a repeated fashion formed the backbone of the DNA molecule (36). A phosphate group is attached to the sugar ring in the position 5′ carbon instead of the –OH. Then, the nitrogen bases are attached to the 1′ carbon of the sugar ring (37). The bases of the DNA adenine (**A**) and guanine (**G**) are purines containing a double ring structure, and cytosine (**C**) and thymine (**T**) are pyrimidines containing a single ring structure (33). A nucleotide is formed by the sugar pentose, the group phosphate and a nitrogen base. A DNA strand is formed by a long chain of nucleotides or polynucleotides joined together. In the molecule of DNA, adenine (**A**) pairs with thymine (**T**) forming 2 hydrogen bonds, and guanine (**G**) pairs with cytosine (**C**) forming 3 hydrogen bonds (33). These effective hydrogen bonds are holding the two chains together (33, 38). The DNA chains run in antiparallel or opposite direction (33, 39). One of the chains of DNA runs 5′ to 3′ and the other from 3′ to 5′ labelled from top to end. The 3′ and 5′ notation come from position of the carbon atom of the sugar pentose in this case deoxyribose (**D**). Then, the 5′ end is the end that has a phosphate group (**P**) attached to the 5 ′carbon. And the 3′ end is the end that has a phosphate group attached to the 3′ carbon or a –OH group if that is the very end of the chain.

This 5′ to 3′ or read it “5-prime to 3-prime” notation is very important in genetic code and genes, and genetic code is always written from 5′ to 3′ direction (37) (Figure 1.2). The sequence of bases in DNA carries the genetic code and particularly and very important sequences are genes. While Eukaryotes keep their DNA in the nucleus, prokaryotes like bacteria have their genetic material DNA, RNA and proteins without showing any organized internal structure inside their plasma membrane (40).

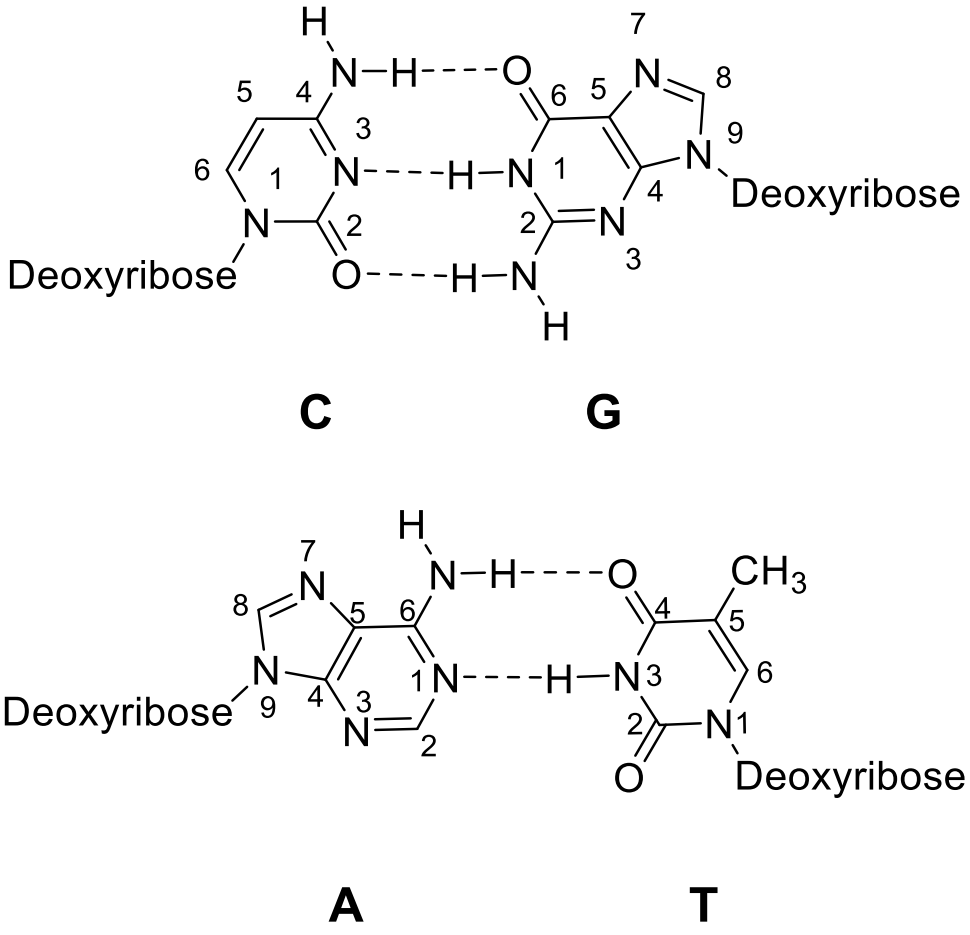


Figure 1.1. Watson and Crick base pairing: cytosine (C), guanine (G), Adenine (A) and Thymine (T).

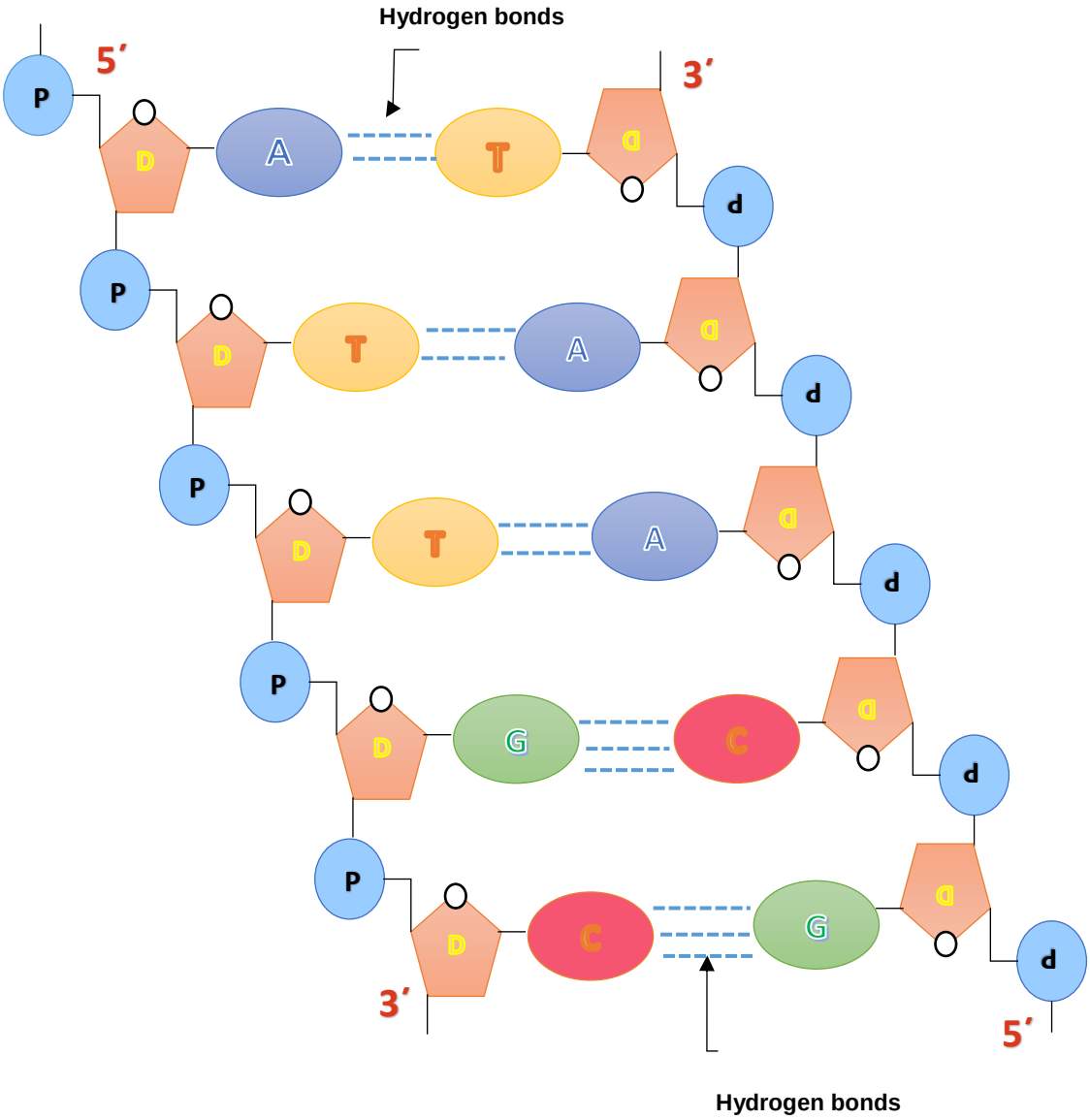


Figure 1.2 DNA structure showing the antiparallel direction of the chains (41). Figure modified and adapted by the author of this thesis.

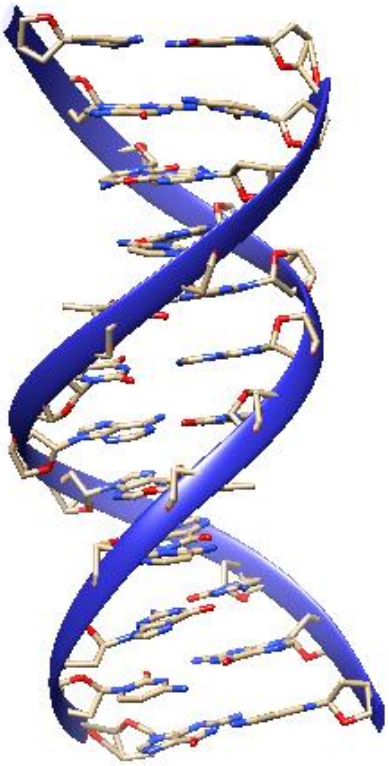


Figure 1.3 B-form DNA double helical structure PDB ID: 1BNA (42).

1.2 RNA

RNA is the information-transfer intermediary between DNA and proteins. Early investigations found that genetic information could not transfer directly from DNA that is located in the nucleus to proteins for the synthesis in the cytoplasm in eukaryotic cell (43). Although RNA occurs most widely in single-stranded form (44), as a consequence is more flexible, and also is capable of folding itself to form double stranded areas with its complementary base-pairing from a single and three nucleotide bulges to a hairpin loop (Figure 1.6). RNA is synthesized from the molecule of DNA mediated by the enzyme RNA polymerase through a process called transcription (45). Therefore, the molecule of RNA is similar to the DNA molecule, but they are some differences in the chemical structure. RNA contains the same nitrogenous bases as in DNA, adenine (**A**), guanine (**G**) cytosine (**C**), but the thymine (**T**) is replaced by the base uracil (**U**) (44, 46). The bases thymine and uracil have very similar chemical structure, but thymine (**T**) in the DNA has a

methyl group (CH_3) in the position 5 of the ring and uracil (**U**) in the RNA does not, and uracil can form the same hydrogen bonds with adenine like thymine in DNA (Figure 1.1 and 1.4). Ribose is the sugar in the backbone of RNA or ribonucleic acid while deoxyribose is the sugar in the DNA or deoxyribonucleic acid. Deoxyribose is a ribose that has lost an oxygen atom in the 2′ carbon of the sugar base (34) (Figure 1.5).

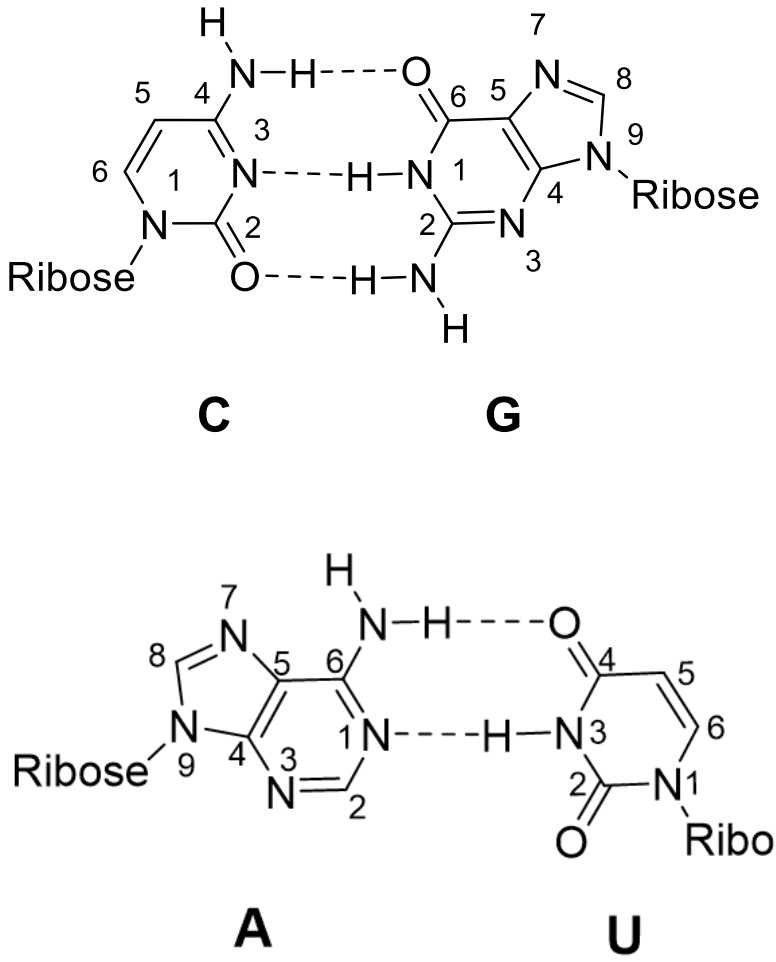


Figure 1.4.RNA base pairing: cytosine (C), guanine (G), Adenine (A) and uracil (U).

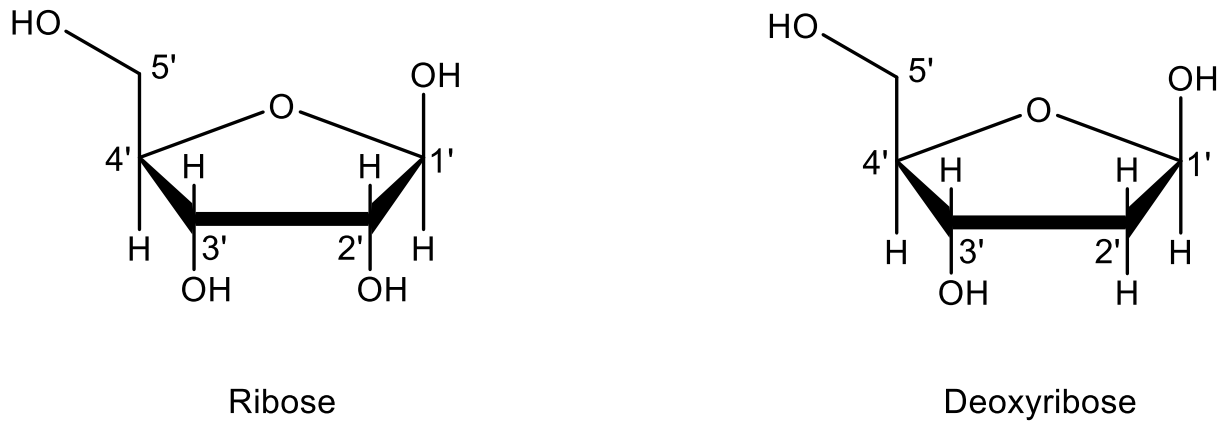


Figure 1.5 Comparing the ribose in the RNA and the deoxyribose in the DNA showing the absence of OH in the carbon 2′.

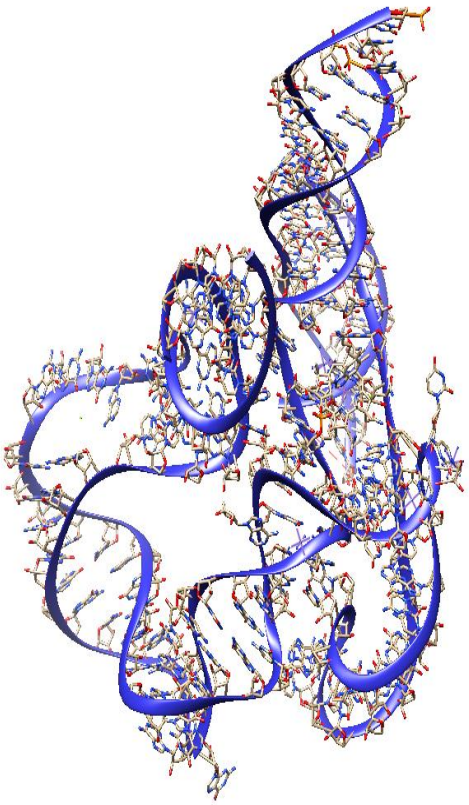


Figure 1.6 RNA structure PDB ID: 4GXY (47).

1.2.1 Central Dogma to Molecular Biology

The central dogma to molecular biology is that genetic information flows from DNA to DNA during replication, from DNA to RNA during transcription, and from RNA to proteins in translation (44, 48). The processes of Replication, Transcription and translation regulate this storage of information and the process of information. This process is used for all living cells eukaryotes (multicellular organisms), prokaryotes (bacteria) and viruses (49) (Figure 1.7).

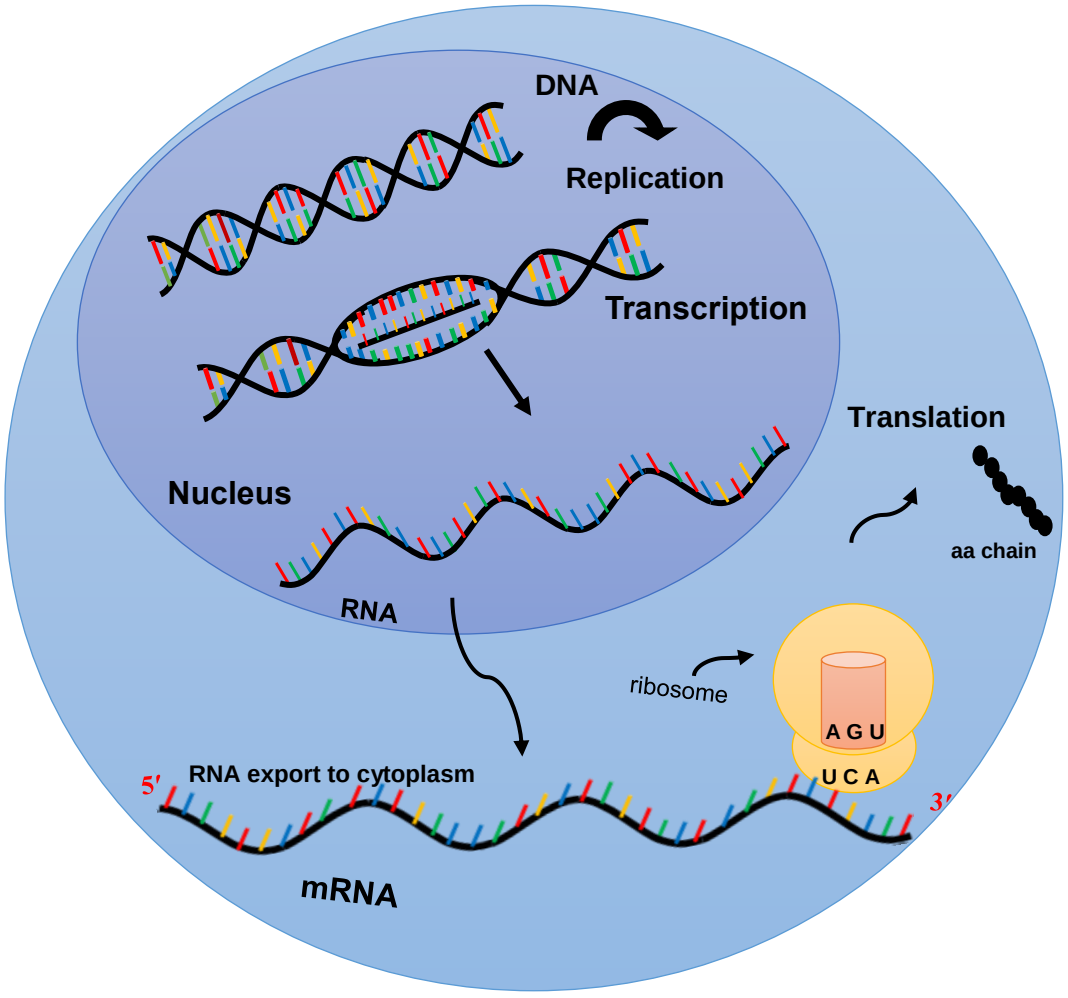


Figure 1.7 Central dogma of gene expression on molecular biology

DNA replication is the process of copying the DNA molecule and require the action of many proteins. This process starts when the DNA is unzipped in two single strands, then, each of the single strands is copied and become half of a new DNA double structure. DNA replication is a semi-conservative process because half of the original DNA and genetic code is conserved in each of the daughter molecules (50). The synthesis of the new complementary strands involves the mediation of the enzyme DNA polymerase which extend the DNA chain only in the 5′ to 3′ direction (51).

Transcription is the process where the genetic information of DNA strand is transferred to the molecule of RNA under control of the enzyme RNA polymerase (52). This process produces messenger RNA (mRNA) which will carry on the genetic code out of the nucleus to the cytoplasm for synthesize of proteins (43). The genetic code is carried only in one of the strands of DNA or the coding strand (43). The coding strand is read from 5′ to 3′ direction. The template strand is complementary to the coding strand, so if the coding strand has a cytosine (C), the template strand has the guanine (G) base pair to match (43). In RNA is the same situation except the uracil (U) is used instead of thymine (T). The transcription process starts when the RNA polymerase finds the start of the gene on the coding strand of DNA or “promoter sequences” that are closer to the 5′ end (43). (Figure1.8 a). Once the RNA polymerase has attached to the DNA, it starts unzipping DNA separating the two strands apart, the coding and template strands, over the length of around 10 bases (Figure 1.8 b). The enzyme starts making RNA by adding new nucleotides at the RNA chain in the direction 3′. Then, the RNA polymerase moves to the DNA and starts zipping it behind closing the space making the RNA new chain become detached from the template strand (43) (Figure1.8 c). Now, messenger RNA (mRNA) was produced and it will be involved directly in the synthesis of proteins (44). Studies about gene expression commonly measure the production of messenger RNA. Most mechanism that control gene expression control transcription or the synthesis of mRNA (53).

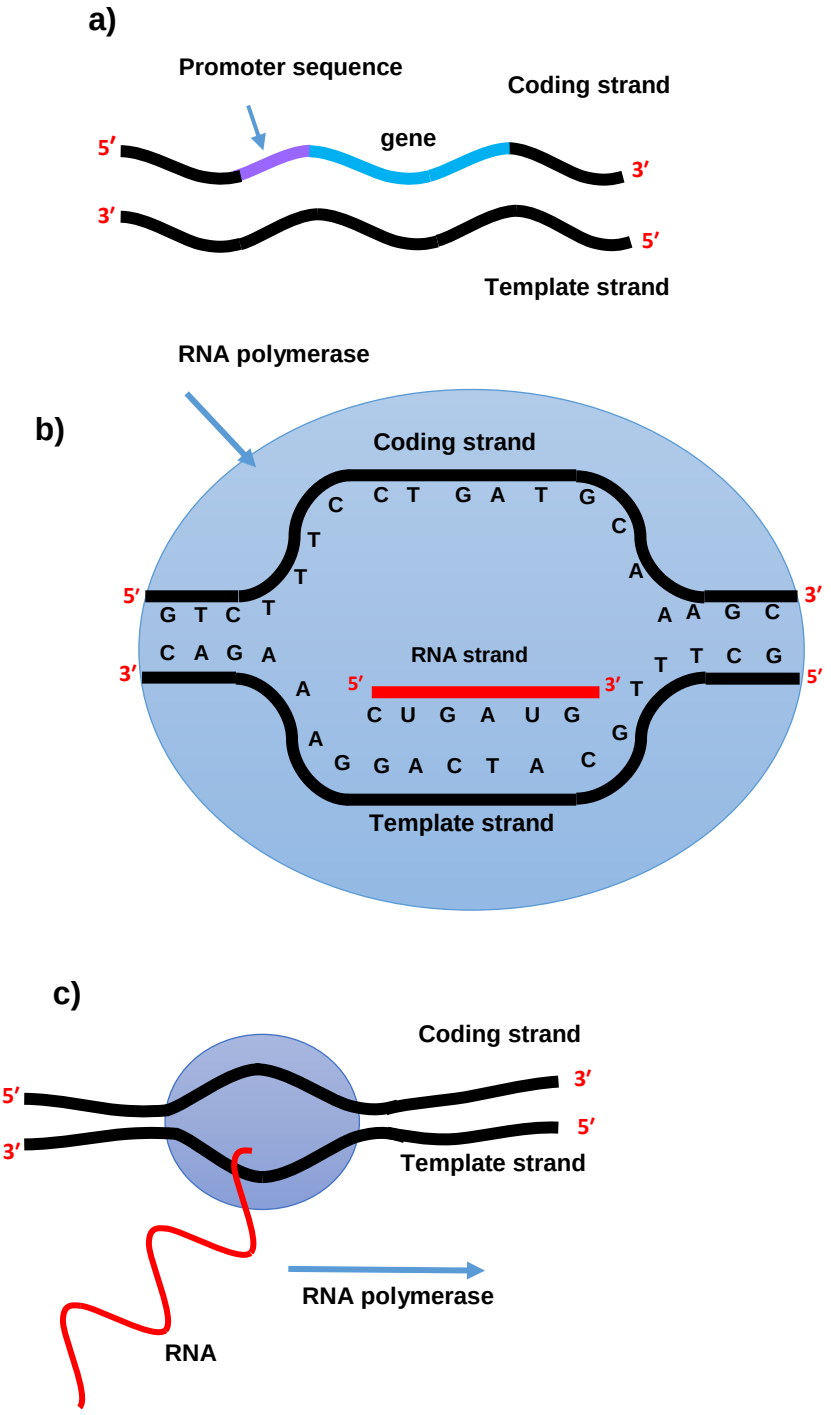


Figure 1.8 Transcription process (41) a) RNA polymerase finding the promoter sequence b) RNA polymerase unzipping DNA strands c) New RNA chain detaching from template strand. Figures modified and adapted by the author of this thesis.

Translation is the process of converting the genetic code of the messenger RNA (mRNA) into a protein chain, and this process takes place in the cytoplasm. The messenger RNA has some sequences of bases or code. An mRNA sequence is decoded in sets of three nucleotides known as codon. In this process of translating the code to build proteins the code has to be carried to the messenger RNA (mRNA) for the transfer RNA (tRNA), all of this is controlled by ribosomes (smaller and large), and other type of RNA ribosomal RNA or rRNA (54). The rules how the nucleotide sequence of a gene is translated into an amino acid sequence of a protein under mRNA are called the genetic code (46). The smaller ribosome starts by finding the start point in the mRNA by attaching itself to the 5′ end to the mRNA and starts moving along until finds a right codon in the mRNA (Figure 1.9 a). Then, the ribosome can start building the protein chain, now the transfer RNA (tRNA) carry on the amino acids to the mRNA and keep them to join together. The tranfer RNA (tRNA) has an anticodon that attaches with the amino acids in the messenger RNA (methionine), then the large risome attaches itself (55) (Figure 1.9 b). Then, another tranfer RNA attaches to the codon of the messenger RNA by its anticodon . The ribosome moves to the next codon in the messenger RNA (mRNA), and a peptide bond is formed between the two aminoacids (methyionine and glycine) (41) (Figure 1.9 c). The process continue until the ribosome reaches a stop codon that does not code for aminoacids, then the chain build up is released from the ribosome.

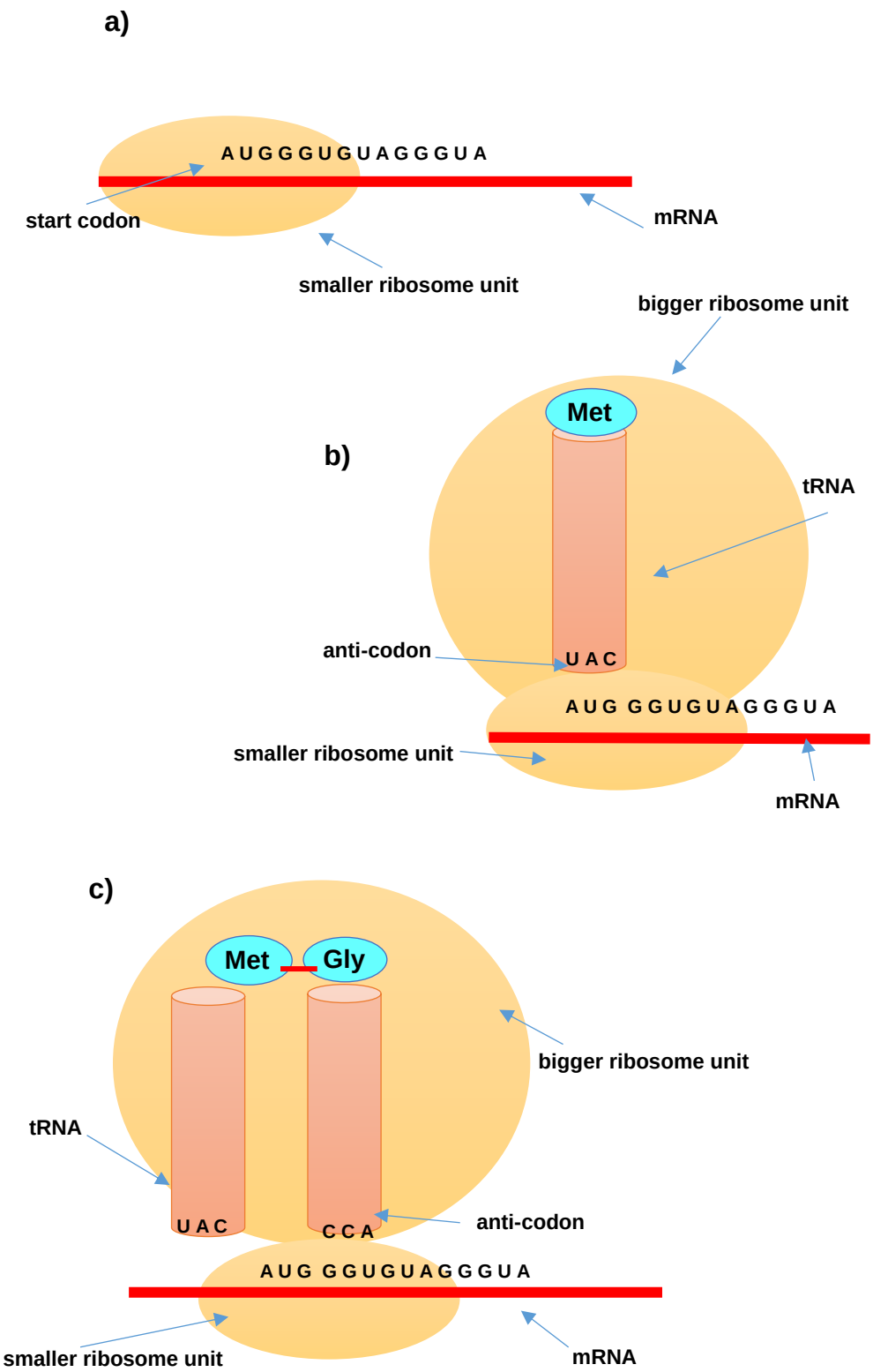
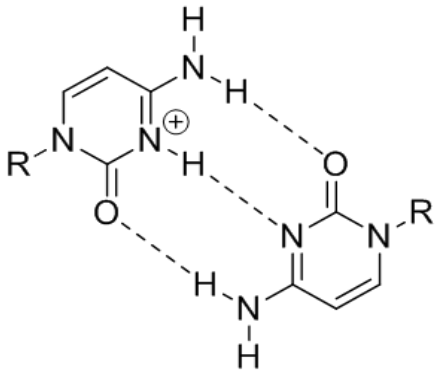


Figure1.9 Translation process (41) a) smaller ribosome unit attaching to the mRNA b) tRNA attaching its anticodon to the mRNA c) Ribosome moving to the next codon. Figures modified and adapted by the author of this thesis.

1.3 i-Motif DNA

DNA structure under physiological conditions take the B-DNA form (56), but also DNA can fold and form other secondary structures like hairpins (34), pseudoknots and left-handed structures (57) like Z-DNA (58) under specific conditions. The i-motifs is also named i-DNA or i-tetraplex (10, 59) are secondary structure of the DNA/RNA molecules formed from sequences rich in cytosine (10). The i-motif structure was first described by Gehring, Leroy and Gueron in 1993 based on early studies conducted by Leroy and co-workers on the oligomers of the deoxycytidine (10). They deciphered the structure composed by d(TC_5) by using NMR, and finding one proton per pair of cytidine at the range of 15-16 ppm suggesting the presence of imino protons hydrogen bonding. This new structure was called i- motif, and its building block is formed for a C-C^+ base pair with one neutral C and one protonated C at the N3 (Figure 1.10 a, b) (10). The conformation of i-motif structure is pH dependent, and highly depends of the pH conditions because in order to this structure to be stable, it needs the hemi protonation of the C-C^+ pair, therefore at the beginning it was thought the conditions of the pH environment has to be slightly acidic (11, 17). In fact, *in vitro* studies revealed originally i-motif formation was dependent on low pH, but currently there are increasing numbers of examples showing i-motif can form at neutral (13, 60, 61) and slightly alkaline conditions (16), gene promoters (62), super helical tension in DNA (63) under molecular crowding conditions (64, 65), and in presence of silver cations (66) and copper cations (67). There are other factors that affect their stability like the loop length (14, 15, 68), the types and number of bases sequence (11, 17), and ligands (28). Due to this pH condition driven, and its ability to fold and unfold under the changes of pH, the majority of i-motif research has been development in nanotechnology, for example nanomachines (69, 70), pH sensors (71-73), and logic operators for computers (74, 75).

a)



b)

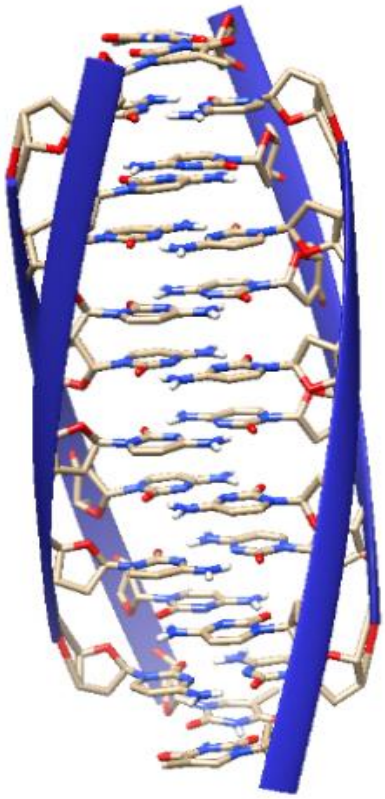


Figure 1.10. The structure of i-motif a) Hemi-protonated cytosine-cytosine pair where R=sugar b) Tetramolecular i-motif structure d(TC_5) PDB ID: 225D (10).

1.4. i-Motif structure

The first i-motif observed was an intermolecular structure formed of four individual strands of cytosine held together through the hemi- protonated cytosine-cytosine pair C-C^+ (10). (Figure 1.11 a) This is the only DNA structure known until now its base pairs have this intercalated spatial arrangement while other structures like G-quadruplexes have their base pairs stacked (3). Also, sequences with multiples C-C+ base pairs aligned in four tracts can fold into an intramolecular i-motif structure (11, 17, 76) (Figure 1.11 b). According to the C-C+ base termination i-motif has been classified in two: the base pair C-C+ at 3′ end (3′E), and with the base pair C-C+ at 5′ end or (5′E) (12, 76). The product of these spatial arrangement produced two major wide grooves and two narrow minor grooves (34, 77). I-Motif structure is naturally unstable because due to repulsion forces cause by the proximity of the phosphates groups charged negatively in the minor grooves of two antiparallel strands. The effect of the phosphate repulsion was investigated using molecular dynamic simulation tests in different topologies of i-motif structure, and they have concluded that either the hydrogen bonds or van der Walls forces or both are the responsible for i-motif stability (78).

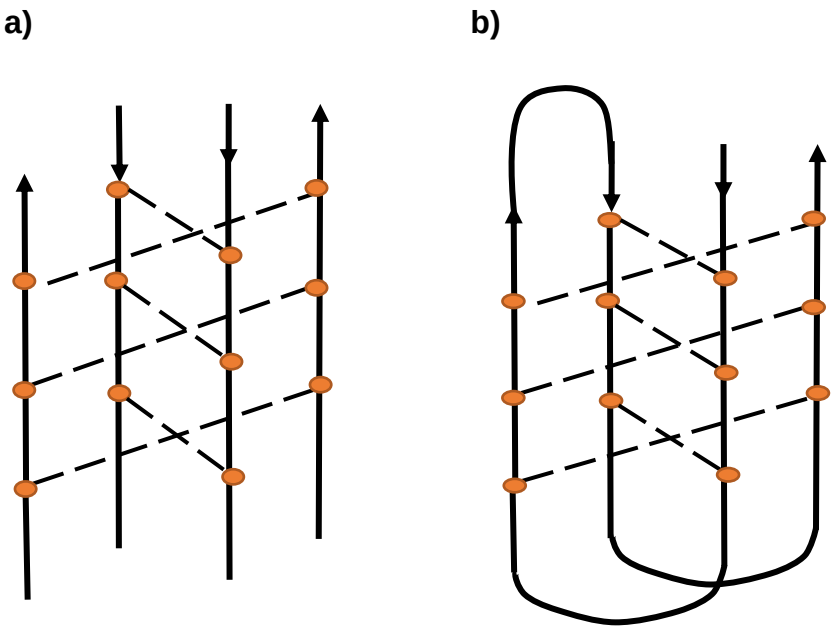


Figure1.11 a) Intermolecular i-motif b) Intramolecular i-motif

1.5 Biological Relevance

Researchers have studied devices at molecular level trying to imitate the complex biological process in the human body considering most of the biological process occurred at nanometer scale (69, 79). i- Motif conformational changes influenced by the pH conditions has gotten the attention because the potential use in nanotechnology to elaborated sensors, nanomachines and pH sensors (69, 79-81). In a study conducted by Wright *et. al.* based on different sequences able to form i-motif under physiological conditions, they found i-motif can form at physiological pH without using any crowding agent, also, the thermal stability increased as the length of the cytosine tracts increased (13). These recent findings might suggest i-motif could form from the numerous amount of cytosine- rich strands in the human genome under physiological conditions. Recently, the existence of stable i-motif in the nuclei of human cells from cytosine-rich sequences in the human genome has been suggested (22, 82). Since sequences rich in guanine form structures called G-quadruplexes which have been found in the human telomeric regions (84, 85), oncogene-promoter regions (8, 18, 85) like bcl-2 (86), c-myc (87), c-kit (88), and their function in the gene transcription has been proposed, we might expect to see the complementary strands rich in cytosine forming i-motif in these oncogene promoter regions, and also, they role in transcriptional regulation has been suggested (63, 20). Hurley's group have observed the effect of the negative supercoiling when duplex strands unwound into single strands, and these single strands might form i-motif structures. This negative supercoiling tension might help the formation of single strands in unwounded DNA regions (63).

1.6 i-Motif RNA

RNA is unstable because the 2′-OH group under slightly basic conditions become deprotonated causing a cleavage of the strand. RNA i-motif are less stable than DNA because the 2′-OH of the ribose produce a twist of the backbone at the minor groove due to steric interactions, and it seems to be responsible for the lack of stability of i-motif structure (23, 89). In general RNA i-motif are less stable than their DNA counterparts. Gehring *et. al*. conducted a UV- melting study that allowed

them to compare the thermal stability of DNA sequences and its RNA counterpart. One DNA residue was replaced by an RNA, and they reported a decrease of the thermal stability (T_m) by almost 7°C. The average melting point for all the DNA i-motif was between 56°C and 48°C, and adding more RNA residues the effects observed were stronger. When two RNA residues were added to the tetrad the melting temperature (T_m) decreased to 19°C affecting the stability of the structure (89). RNA demonstrated a preference to adopt a triplex structure than i-motif structure (23). It was observed when uracil was substituted did not affect the stability of i-motif, in consequence, the methyl group of thymine does not a key part on the stability of i-motif, then these findings suggested it must be the 2′-OH group in RNA (23). According to Collins *et. al.* the steric hindrance between the 2′-OH in the narrow grooves might cause of the inability of RNA to form i-motif structure (89). Damha's group also have studied the effects of the stability of i-motif structure by characterizing the sugar interactions, and observed a weak or not distinguishable UV melting temperature for the RNA counterpart suggesting a lack of stability (90).

1.7 Techniques to investigate i-Motif

Different biophysical techniques have been used to study the formation and stability of i-motif structure.

1.71. Nuclear Magnetic Resonance (NMR) spectroscopy

NMR is a biophysical method used to determine and analyze the molecular conformation in chemistry, physical properties at molecular level in biology, screening drugs, and their products in pharmaceutical field by detecting the orientation of nuclear spins in an applied magnetic field (91). Gueron and co-workers resolved the i-motif structure in 1993 in the DNA sequence 5′-d(TCCCCC) at acid pH using NMR technique, and the structure revealed a four-stranded system formed with two base pair duplexes oriented in antiparallel direction with their base pairs intercalated. NMR spectrum revealed important information about the formation of the C-C^+ pair showing a six-spin system which present a symmetrical structure having the four strands symmetrical (10). The C-C^+ base pair showed an NMR signal

at 15 -16 ppm (10) while the base pairs A-T and C-G of Watson and Crick pairing show an NMR signal at 12-14 ppm (17, 92).

1.72 X-ray diffraction (XRD)

XRD technique was described as a new high-speed technique based on X-ray spectrography, and as a new method applicable to crystal structures where the samples are irradiated with polychromatic radiation from an x-ray tube, then the energy spectrum of x-rays scattered at a given angle is observed with a semiconductor radiation detector coupled with a pulse-height analyser (93). Over 30 years before i-motif structure was resolved in 1993 using NMR technique (10), Marsh *et. al*. in 1962 have observed in the crystal of cytosine-5-acetic acid that the base cytosine can form a hydrogen-bonded pair with itself when the cytosines bases were hemi protonated (94). Shortly after, Langridge and co-workers in 1963 have found after analysing the poly-cytidylic acid using XRD technique that two parallel strands were held together by this type of hydrogen bonding and stabilized only when the pH was below 7 (95).

1.73 FRET (fluorescence resonance energy transfer)

FRET technique has been used more and more in medicine research and drug discovery. FRET efficiency is distance-depend transfer of the excited state energy between a fluorescent donor to an unexcited acceptor (96). FRET can be used to measure nanometre scale distances and the changes in distances, both *in vitro* and *in vivo*. Mergny *et. al.* used this technique to analyse the folding of i-motif using fluorescent probes showing that the intramolecular folding of i-motif generates state of fluorescence excitation between a donor an acceptor covalently attached to the 5′ and 3′ ends. These results allowed the study of the thermal stability of i-motif structures since the folding or unfolding of DNA was observed by measuring the change in fluorescence (97). FRET

screening assay was used to identify a group of BCL2 i-motif interactive compounds that stabilised or either destabilised i-motif structure demonstrated by the change of fluorescence intensity by Hurley *et. al.* (20). FRET technique has been used to monitor and measure the melting temperature (*Tm*) of i-motif structure. Mergny and co-workers showed that monitoring the fluorescence with increasing temperature results in a sigmoidal curve from which the melting temperature can be identified in a similar way to UV melting curves (98). FRET technique is only showing that folding is occurring in the i-motif structure since the two fluorophores are close together, but to prove that i-motif formation is occurring a different technique is necessary such as ultraviolet spectroscopy (UV) or circular dichroism (CD).

1.74 X-ray crystallography (XRC)

XRC technique have been used to determine the three-dimensional molecular structure from a crystal. After the crystal is exposed to X- ray beam causing diffraction patterns. The resulting diffraction patterns can then be processed providing information about the crystal symmetry and size of the crystal units (99). XRC was used to solve the crystal structure of C_4 at 2.3 Å resolution revealing the origination of a four-stranded molecule formed of two intercalated duplexes (100). XRC provides different information from solution-based techniques such as NMR. The atomic interactions within the molecule dominate in solution state. But, in the crystalline state, XRC can provide information about the packing interactions formed between molecules. Berger and co-workers have observed that the crystal structures of d(CCCT), d(TAACCC), d(CCCAAT), and d(AACCCC), and these molecules have showed in common intercalated cytosine segments that have similar geometry (101).

1.75 PAGE (polyacrylamide gel electrophoresis)

PAGE technique used to separate and distinguish between DNA, RNA and proteins fragments according to their size based on their mobility

through the gel (102). DNA fragments are negatively charged due to all the phosphate groups in the backbone of DNA, so they move towards the positive electrode when an electric current is applied to pull them through the gel. This property is called electrophoretic mobility shift assay or EMSA (103). It depends of the molecular weight of the DNA complex to move slowly or faster through the gel, and by measuring distances and comparing the position of stain to the control sample we can determine if any change has occurred. In order to compare results, it is important to have the same the same control sample in one of the lanes in each of the gels under the same conditions. PAGE technique has been used to demonstrate the formation of i-motif structure. For example, Zhou *et. al.* have shown i-motif structure can form at neutral and slightly alkaline pH at 4°C (16). And, Simonsson and co-workers have observed the i-motif formation can occur in the promoter region of the human c-myc gene *in vitro* suggesting the biological role for i-motif structures *in vivo* (87).

1.76 Ultraviolet (UV) spectroscopy

UV spectroscopy is a very useful method to stablish the thermal stability of i-motif structure. DNA can absorb light at 260 nm, and folded DNA absorbs light different than unfolded DNA due to the π-π stacking interactions. With UV spectroscopy, we can observe changes in i-motif structure due the protonation of the cytosines which produce an increase (hyperchromity) at 260 nm wavelength, and a decrease (hypochromism) of the absorbance at 290 nm indicating fold into the structure (98). For instance, changes in the absorbance around 295 nm for i-motif with increase of the temperature or melting temperature (T_m) provides information about the stability of i-motif structure. UV spectroscopy have been used to observe the stability of i-motif structure in function of temperature and pH. Mergny *et. al.* have found a stable i-motif structure at pH values closer to the pKa value of

cytosine that is around 4.67 at 25°C, and at pH values lower than 6.6. When pH values were higher, the C bases become deprotonated and the structure unfolds into a single strand. On the other hand, at lower pH values below 3 the C-C^+ base pair was not able to form because all the C bases become protonated and they cannot form the hydrogen bond that is necessary for the formation of the C-C^+ base pair (11). Wright and co-workers have identified a library of i-motif forming sequences with cytosine tract varying at pH 7.4 using UV spectroscopy among other techniques (13).

By heating the sample, we can record the melting temperature (T_m) of the sample or the temperature of transition when the structure change from the folded to the unfolded stage (Figure1.12 a) (104). The first derivative method is used to calculate the melting temperatures (T_m). First derivative is calculated by dividing the change of absorbance and the change of the temperature (ΔAbs/ΔT°) for every cycle (see Materials and Methods) (Figure 1.12 b). And, by comparing melting temperatures of different i-motif forming sequences, we can compare and evaluate stabilities. Brazier and co-workers have found stable i-motifs sequences near neutral pH complimentary to already known G-quadruplexes (62)

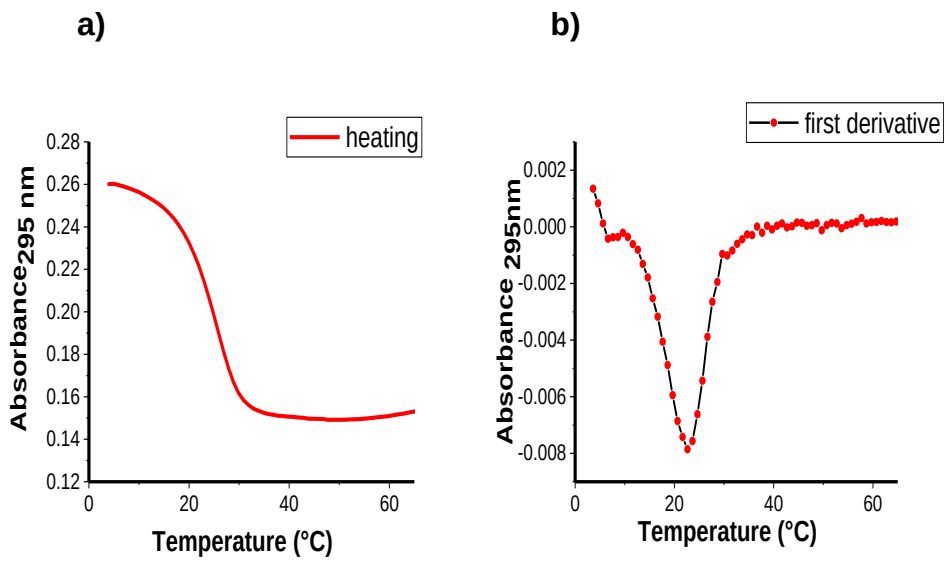


Figure 1.12 Example of a) melting curve and b) a first derivative for the heating cycle1 i-motif at 2.5µM RNA in sodium cacodylate buffer (10 mM pH 5.5) and NaCl (100 mM).

By monitoring the changes in the absorbance at specific wavelength and subtracting the UV spectra of the folded (low temperature) from the spectra of the unfolded (high temperature) a thermal difference spectra (TDS) is calculated (10). TDS is very valuable tool for identifying the specific type of secondary structure present in the sample because is very unique, and has specific shape for every nucleic acid reflecting the base their stacking interactions. The thermal difference spectra of i-motif DNA structure showed a major positive peak around the 240 nm and a negative peak around 295 nm (105). In the figure below the two curves are reflecting the arithmetic difference of the spectra between the high temperature and low temperature. In order to facilitate the comparison of the spectral curves, the difference spectra was normalized by dividing our raw data by its maximum value as a result we get Y value of +1 corresponding to the highest positive peak (105) (Figure 1.13).

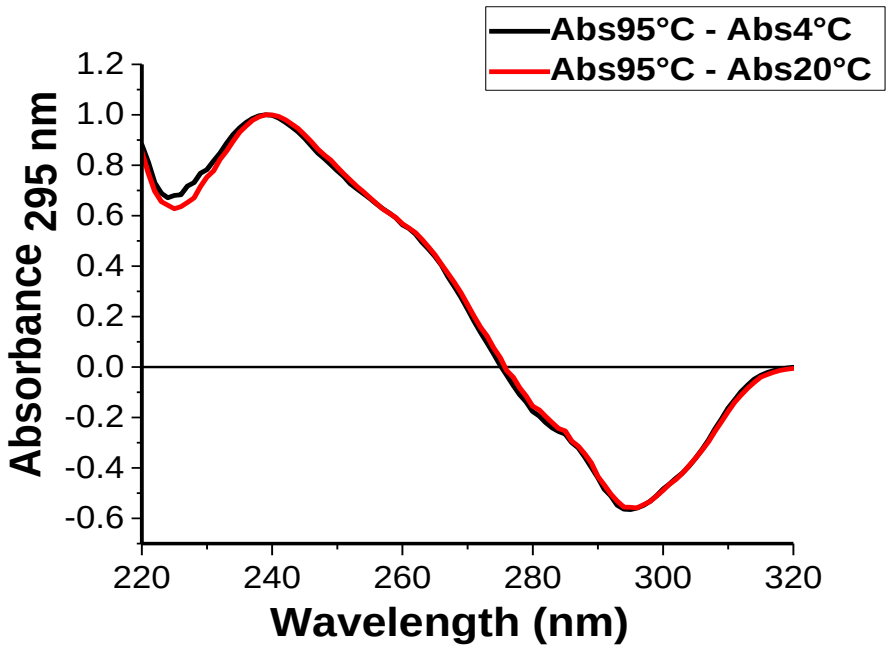


Figure 1.13 Thermal difference spectra for i-motif DNA structure at 2.5µM DNA in sodium cacodylate buffer (10 mM pH 5.5) and NaCl (100 mM). Data obtained from Wright *et. al.* (13).

1.7.7 Circular Dichroism spectroscopy (CD)

Circular dichroism is a phenomenon produced by the interactions of circularly polarized electromagnetic rays and chiral molecules (106). DNA secondary structures have different optical activities and therefore reveal different circular dichroism due to the different of their glycosidic bond angles has given optical activities and circular dichroism. CD spectroscopy is a fast, simple and economical method to confirm the formation of i-motif structure, and provide information about the conformational properties of the molecule of DNA. This method can be used with a low or high concentration of DNA, also it can be used with short or long DNA molecules. Moreover, the samples to be analyzed can be titrated with different substances to see any conformational changes of the molecule of DNA (105). By measuring the CD signal of

a sample in function of wavelength change provides a specific spectrum on the type of DNA present in the sample. DNA i-motif structure poses a characteristics CD spectrum or two bands one positive and dominant band at 288 nm and one negative at 260 nm (105). (Figure 1.14).

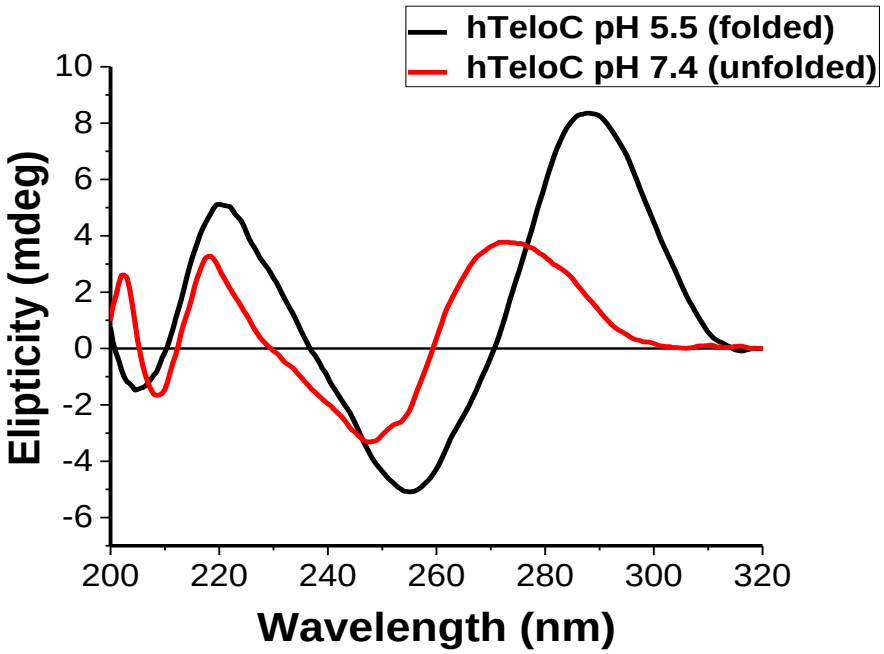


Figure 1.14 Circular dichroism of human telomeric DNA i-motif (10 μM) in sodium cacodylate buffer (10 mM) and NaCl (100 mM). Data obtained from Wright *et. al. (13).*

1.8 i-Motif in plants

Plants and animals have different responses to stressful changes in their habitat. Plants and animals under some stress conditions adopt some mechanism to tolerate, cope, mitigate and then, recover. While animals can use movement to alleviate stress, plants cannot move like animals (106). Plants develop other mechanism at molecular level, and physiological mechanism to tolerate stress (26). For instance, plants have higher tolerance to changes on their habitat like pH (107), drought conditions (108), and dramatic changes in temperature (109) and increased of salinity (110). The lack of water and high concentrations of salt are the main cause for the reduction of growth in crops worldwide (108), also they affect plants by diminishing their photosynthesis and ability to absorb of nutrients (110). The pH is an important factor for the plant cells, and, pH is highly variable in soil affected by ionic and organic composition, and water (111, 112). It has observed low levels of pH in acidic soils can cause poor plants development, and also can cause the efflux of K^+ depolarizing the plasmatic membrane (30, 107). Plants can modify apoplastic pH in response to external signals. Inside plants, the apoplast is the space outside of the plasmatic membrane where material can move freely.

Even though, they are a considerable amount of studies of plant and gene expression (113, 114), studies about how the pH affects regulation on gene expression has not been widely reported. Larger *et. al.* have observed the effect of external pH in gene expression in a study conducted in *Arabidopsis thaliana* roots, 881 genes showed at least two fold changes in transcription 8 hours after the pH was changed from 6 to 4.5 (32). On the other hand, *Gao et. al.* found the external pH of the roots of *Arabidopsis thaliana* have a strong affect in the apoplastic pH, but not in the cytosolic pH (115). From our knowledge, there are not studies in the literature focused in the

mechanism of post-transcriptional (between the transcription and translation process) regulation at RNA level and i-motif in plants.

1.9 i-Motif and molecular crowding

In order to understand the complex machinery of the structure of biomolecules *in vivo*, researchers perform studies *in vitro* trying to predict their properties. The intracellular medium is crowded with different macromolecules counting around 20-40% of the total of the cell volume (116-118) under normal conditions both for animal and plants, however, it could reach 90% under severe desiccation in plants (119). Nevertheless, most of the studies are performed in attenuated conditions of the environment (117, 118). Studies of how crowded conditions affect the stability and structure of macromolecules DNA/RNA are very important tool these days. A crowding agent can mimic crowding conditions inside the cell environment (118). Among the principal characteristics of crowding cosolutes: it should be neutral, not able to produce any interaction in the molecule of interest, soluble in water to activate the crowding conditions, and available in different sizes (120). The most common cosolutes are PEG (polyethylene glycol), ficoll and dextran (121). Hemoglobin, albumin and lysozyme are proteins that have been used to mimic crowding conditions (120). Macromolecular crowding conditions stabilized proteins in their folded state (65, 122). Also, it can accelerate the folding process because a compact folded protein will occupy less volume or become more compact than an unfolded chain (123).

i-Motif structure has been observed at neutral and slightly basic pH under the presence of molecular crowders. For example, Sugimoto and coworkers have observed the presence of DNA i-motif at 20% PEG 200 and 8,000 at pH 6.6, 7.1 and 7.4 (64). One of the effects of molecular crowding inside the cells is called excluded volume. The

excluded effect is the volume occupied by a molecule and not accessible to other macromolecules (117). For example, Kilburn *et. al.* observed that the excluded volume effect has a higher influence on the folding process with a molecular crowder like PEG than the disturbance caused by the activities of ions or water (122). They have found the PEG 1,000 lowered the midpoints of the folding transitions by stabilizing the folded stage of RNA. They have found that only 0.016% of the ribozymes were folded in the absence of PEG 1,000, and 50% folded in the present of 20% PEG 1,000 (122). In addition, Strulson *et. al*. worked with three different sizes of PEGs: 200, 4000 and 8000, and observed all of them promote RNA folded cooperativity in presence of low biological concentration (0.5-2 mM) of Mg^{+2}. They concluded the influences of the crowder are the result of the excluded volume effect (123).

Mimicking crowding conditions in cells can increase the stability of i-motif at acid pH and cause dehydration (124). Dehydration can affect the folding activity of DNA (120). Moreover, there are changes in the physical properties of the solution as a consequence of the molecular crowding. By adding molecular crowders to a solution, the viscosity of the conditions are affected as well the rates of the chemical reactions (125). As a consequence, the viscosity not only affects the rate of diffusion inside the cell, but also the folding process of biomolecules (126).

The Bhavsar group also suggest the stability i-motif has been influenced by not only molecular crowding but also dehydration (127). They found cosolutes can be modified DNA secondary structure by dehydration effect. They quantified the released of water molecules related with the DNA structure at 37°C in pH 5.4, and the results indicate that in presence of glycerol they were 6 water molecules related to the folded DNA while ethylene glycol released 1.4 water molecules per i-motif structure, and the PEG 300 released 26 water

molecules. Their findings suggest molecular crowders stabilize i-motif by reducing the hydrogen bonding through water molecules. Additionally, they tested PEG 300 from 10-30% at different pH, and it was found the PEG 300 stabilized i-motif versus both temperature and pH by releasing of water molecules during the folded process at higher pH (127). Apparently, the ligand affinity and stabilization increased with a higher mass of PEG (128). Lewis's group observed by adding a large molecular crowder like PEG 8,000 at 30% and 40% seems to be responsible for stabilizing c-myc i-motif structure at values close to physiological pH (65). Also, It has been reported an increase in i-motif melting temperature with cosolutes with higher molecular weight (120). For example, Rajendran *et. al.* have observed the melting temperature i-motif is higher at PEG 8,000 (64).

1.10 Aims and objectives

The aim of this project is to investigate the stability of RNA i-motif. This work will be carried through in 2 stages described in this dissertation. The first part in chapter 2, ten cytosine-rich i-motif forming RNA sequences, version of the DNA sequence $[C_nU_3]_3C_n$, are investigated to determine the presence and thermal stability of RNA i-motif at pH 5.5 and biological pH 7.4 using some biophysical techniques like Ultraviolet spectroscopy (UV), circular dichroism (CD) and thermal difference spectra (TDS). This is in order to confirm this already studied DNA i-motif stable sequence at pH 7.0 can fold into a RNA i-motif structure. The second part in chapter 3, a plant sequence AT5G08230.1 RNA found in *Arabidopsis thaliana* is characterized at pH 5.8. The presence of RNA i-motif structure is going to be determined by using circular dichroism (CD), the thermal stability of RNA i-motif using ultraviolet spectroscopy (UV). The plant sequence AT5G08230.1 RNA is going to be modeled under crowding and

drought stress conditions. PEG (polyethylene glycol) 8,000 at two different percentages 20% and 40% is going to be used as a molecular crowder to mimic the crowding conditions in cells. Drought conditions to see the K^+ influx by using KCl at 100 mM and 800 mM. The purpose of these is to confirm RNA i-motif structure can form in plants at acidic pH where plants normally thrive. Also, this is in order to see a potential role of RNA i-motifs structures in plants to help dealing with stress due to i-motif on/off switch functions driven by pH conditions.

CHAPTER 2: Studying the effect of C-tract length on stability of RNA i-Motifs using the sequence C_nU_3.

2.1. Introduction

Our principal aim was to characterise RNA cytosine-rich forming sequences at neutral pH by establishing the thermal stability, and the effect of varying the cytosine tract length on the stability of i-motif structure for all ten sequences from our library (Table1). Gehring *et. al.* have demonstrated DNA can form i-motif structure which are stable under acidic conditions (10), it was originally believed that i-motif formation was able to form only on low pH, but they are currently studies showing i-motif can form at neutral and slightly alkaline conditions (13). The RNA i-motif structure has not been widely studied, and the literature around this theme is very limited due to the previous thought to their low stability. Lacroix and co-workers showed RNA i-motifs are less stable than their DNA counterparts due to a clash in the minor groove caused by the 2'-OH groups in the sugar ribose (23). Previous work in Waller's Lab has have systematically studied a cytosine rich DNA sequence which tract length of consecutive cytosines has been modified in order to investigate their structure stability and potential capacity to fold and form i-motif structure under physiological conditions (13). They have demonstrated using a library of sequences that the sequences with 5 cytosine per tracts were capable of folding into i-motif at neutral pH without using a crowding agent (13). In the light of this evidence, we hypothesise that it might be possible for certain RNA sequences to be stable at neutral pH. So, we have studied the RNA equivalent of the same library of DNA sequences to compare: $[C_nU_3]_3C_n$ where n goes from 1 to 10. We have characterized the thermal profile of all these cytosine-rich RNA sequences form our library (Table1) using some biophysical techniques such as Ultraviolet (UV) spectroscopy, Circular dichroism (CD) and thermal difference spectra (TDS). We have used sodium cacodylate as a buffer because does not absorb light in the far UV range of 180-320 nm, and it does not vary with the temperature, and additionally work well at pH 5.5 and 7.4 (104). After we have established the presence of a secondary structure, thermal difference spectra (TDS) was used to determine the specific secondary structure present in every sequence of our library, and additionally circular dichroism (CD) was used to further characterise the secondary structures related to the melting profiles.

2.2 Thermal profile studies with UV spectroscopy at pH 5.5.

2.2.1 Introduction

We have studied and characterized all the sequences from our library at pH 5.5 since we have known from the literature i-motif formation is stable under acidic condition (10). UV spectroscopy was used to determine the melting (T_m) and annealing (T_a) temperatures for all sequences from our library (Table1).

2.2.2 Results and Discussion

All the sequences from our library have shown a folded and an unfolded state on their melting and annealing curves at pH 5.5 except C_1U_3 (Figure2.1a). Although C_2U_3 cytosine tract length was increased with just one more cytosine, we observed an increase of thermal stability for the melting and annealing curves (Figure2.1b). The three melting (T_m) curves for C_2U_3 have showed an average of 7.3 ± 0.6 °C, and 8 ± 0.9 °C for the annealing (T_a) curves (Table1). At pH 5.5 conditions we have observed an increase in the T_m for all the sequences when the number of cytosines per tract increase (Figure 2.2 **a**). We have found the increase in the temperature was marked from C_2U_3 to C_3U_3 with a change of the T_m of 9°C, after tract lengths of three cytosines the increase in T_m was smaller (Table 1). For example, the increase in T_m from C_3U_3 to C_4U_3 was 2°C, from C_4U_3 to C_5U_3 the increases in T_m was 3°C, C_5U_3 to C_6U_3 was 3°C. After tract lengths of 5 cytosines, the increases in T_m was 3°C until C_9U_3 to $C_{10}U_3$ was 5°C. (Table1 and Appendix A1). For both C_nT_3 and C_nU_3 at pH 5.5, we have found the melting temperatures increase with the number of cytosines per tract increase, but the average melting temperatures for the C_nT_3 was found higher suggesting higher stability than the C_nU_3, for example the $C_{10}T_3$ had a T_m of 67°C while $C_{10}U_3$ had a T_m of 37°C (Figure 2.2 a and Figure 2.1 d).All these sequences from our library showed an increase in the T_a with increasing the number of cytosines per tract but C_1U_3 **(**Figure 2.2 b**)**, we have observed an increase in the T_a from C_2U_3 to C_3U_3 of 7°C (Table1), but after tract length of 3 cytosines the increase was minimal suggesting

the annealing temperatures reached a plateau (Figure 2.2 b and Appendix A1-b, **c**). Then from C_5U_3 to C_6U_3 the T_m increased in 4°C (Figure 2.1 c and Table1), and the increases after seven cytosines were not significant for example C_7U_3 to C_8U_3 a change in T_m of 1.5°C, C_8U_3 to C_9U_3 a change in T_m of 1.4°C (Table1), and there is an increase of the T_m for sequences with longer cytosine tract length. C_9U_3 to $C_{10}U_3$ the T_m increased in 4.7°C (Appendix A1-i, j and Figure 2.1 d). We have found for both C_nT_3 and C_nU_3 the annealing temperatures increase when the number of cytosines per tract has increased (Figure 2.2 b). Overall, the C_nT_3 annealing temperatures are higher than the C_nU_3 confirming the more stable thermal stability of the C_nT_3 in compare with C_nU_3 with the $C_{10}T_3$ having a T_a of 65°C and the $C_{10}U_3$ having a T_a of 24.2°C (Table1 and Figure 2.2 b). The tract length of these sequences could have an effect in the kinetic of dissociation (melting) and association (annealing) according to Mergny *et. al.* (11). Hysteresis between the melting and annealing curves suggested the reactions are not at thermodynamic equilibrium. For all our sequences from our library the degree of hysteresis increased when the cytosine tract length increased. But, this degree of hysteresis reached an upper point in the sequences with cytosine tract length of six, seven or more (Figure 2.2 c). The degree of hysteresis versus tract length was compared for both C_nT_3 and C_nU_3, and Wright *et. al.* have observed the degree of hysteresis reached an upper point for C_nT_3 with sequences of cytosine tract length of six or more also (13), but the degree of hysteresis was higher for the C_nU_3 in general suggesting less thermal stability in compare with the C_nT_3 (Figure 2.2 c)

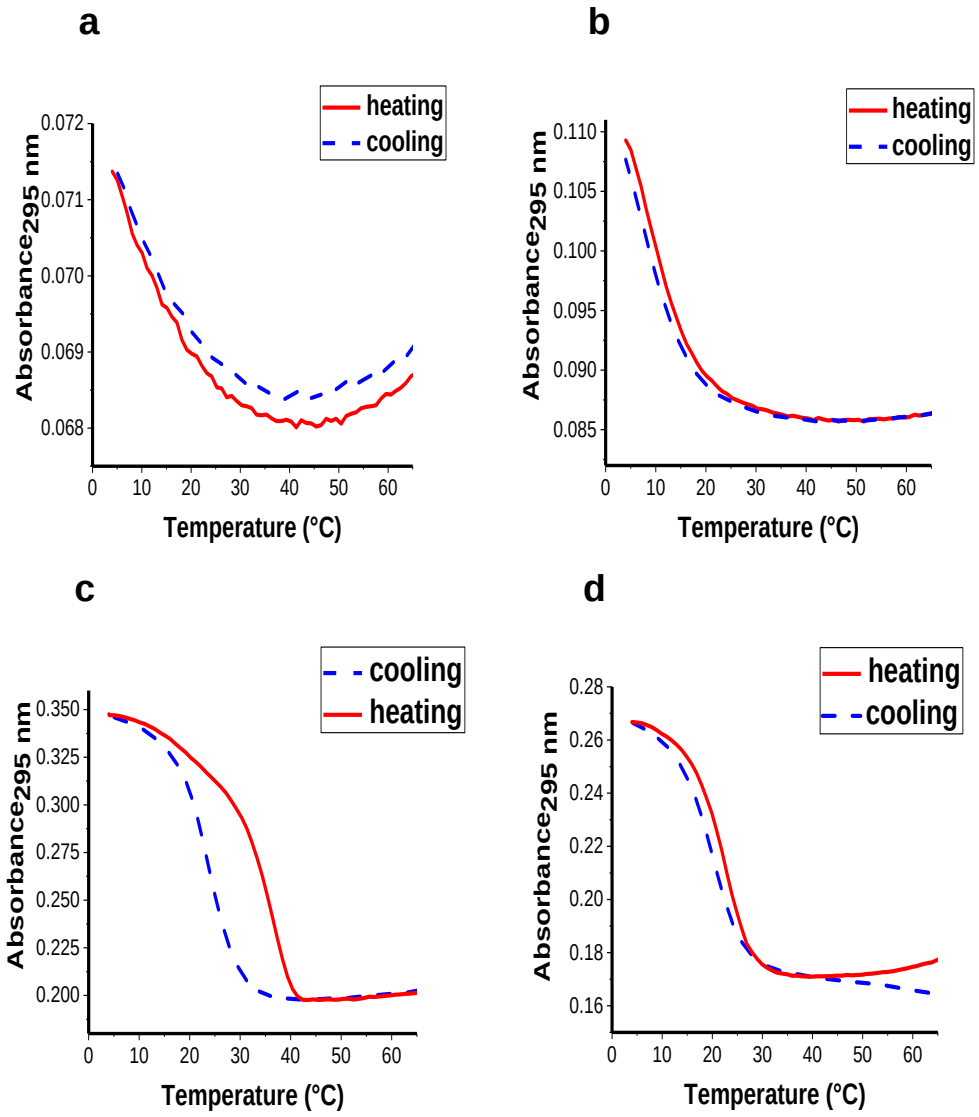


Figure 2.1 UV melting/annealing representative profiles of 2.5µM RNA in sodium cacodylate buffer (10 mM pH 5.5) and NaCl (100 mM) (**a**)C_1U_3, (**b**)C_2U_3, (**c**)C_5U_3 (**d**)$C_{10}U_3$.

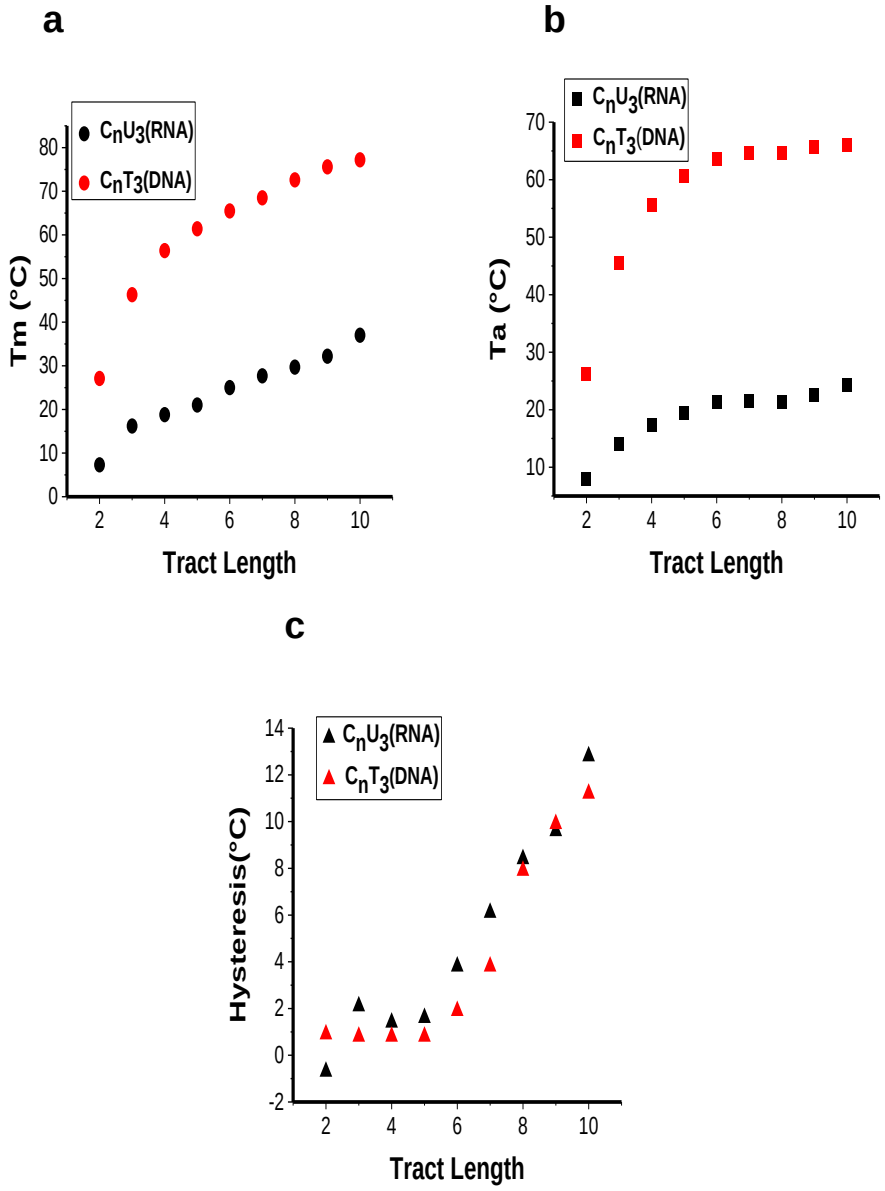


Figure 2.2 Thermal stability vs. cytosine tract length for C_nT_3, C_nU_3 (**a**) Melting temp. (T_m) (**b**) Annealing temp. (T_a) (**c**) Cytosine tract length vs. hysteresis of 2.5µM RNA or DNA in sodium cacodylate buffer (10 mM pH 5.5) and NaCl (100 mM).

After we have confirmed the presence of secondary structures using UV spectroscopy, we have used TDS to stablish the predominant folded structure for all our sequences at pH 5.5 (Appendix A2-a to j). We found all our sequences folded into i-motif structure showing a positive and dominant peak at 240 nm and a negative peak at 295 nm except for the C_1U_3 (Figure 2.3 a). These findings are

compatible with the TDS of i-motif showed previously in Chapter1 obtained from the DNA i-motif forming sequence with varying tract length C_nT_3 (13) (Figure 1.13). Both C_nT_3 and C_nU_3 have showed sequences from 2 to 10 cytosines per tract length folded into i-motif (Figure 2.3 a, b).

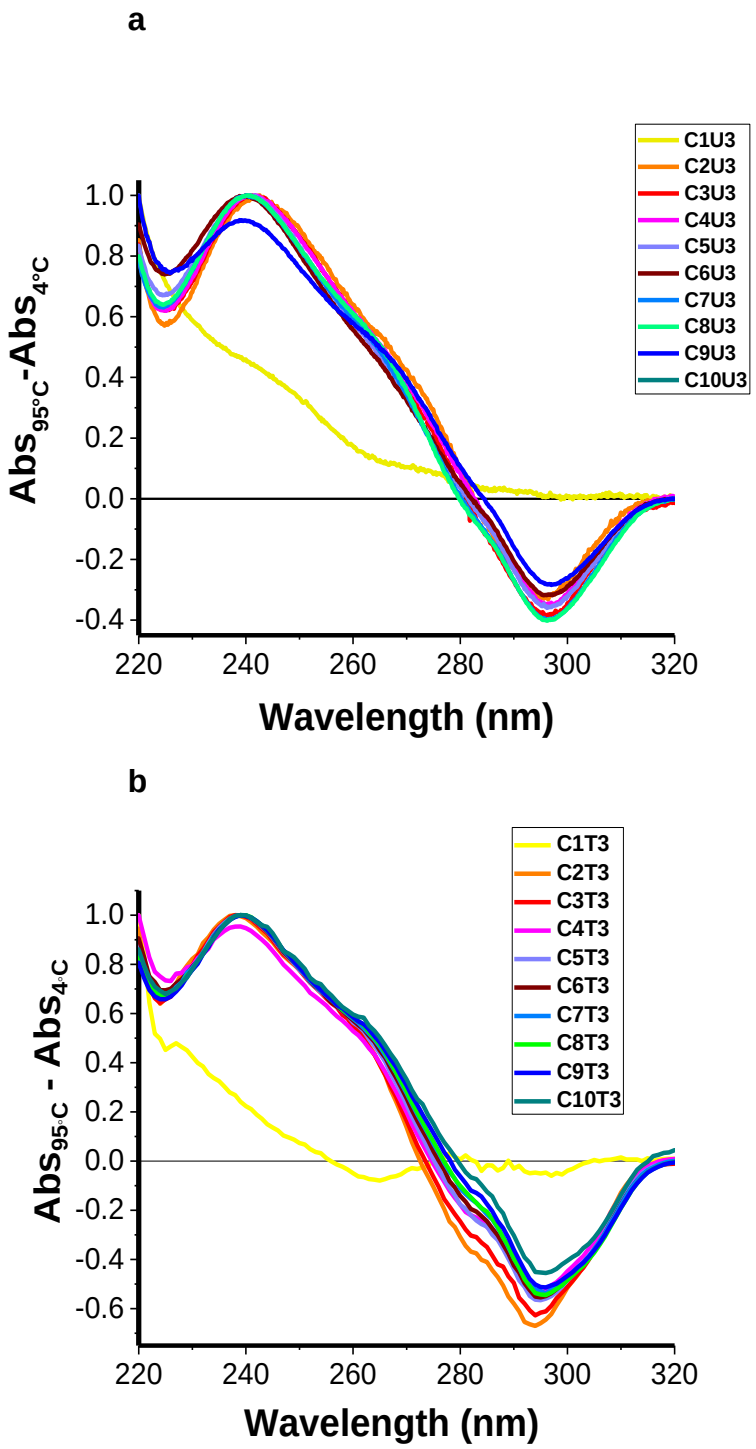


Figure 2.3 The Thermal difference spectra calculated between 95°C and 4°C at pH 5.5 of (**a**) CnU3 (RNA) and (**b**) CnT3 (DNA) of 2.5µM in sodium cacodylate buffer (10 mM pH 5.5) and NaCl (100 mM).

Sequence 5′- 3′	Bases	Notation	pH 5.5	
			T_m	T_a
$[C_1U_3]_3C_1$	13	C_1U_3	nd	nd
$[C_2U_3]_3C_2$	17	C_2U_3	7.3 ± 0.6	8.0 ± 0.9
$[C_3U_3]_3C_3$	21	C_3U_3	16.2 ±0.6	14.1 ± 1.7
$[C_4U_3]_3C_4$	25	C_4U_3	18.8 ± 0.9	17.4 ± 0.6
$[C_5U_3]_3C_5$	29	C_5U_3	21 ± 0.9	19.4 ± 0.6
$[C_6U_3]_3C_6$	33	C_6U_3	25. 1± 0.4	21.3 ± 0.6
$[C_7U_3]_3C_7$	37	C_7U_3	27.7 ± 0.8	21.6 ± 0.6
$[C_8U_3]_3C_8$	41	C_8U_3	29.7 ± 0.7	21.3 ± 0.6
$[C_9U_3]_3C_9$	45	C_9U_3	32.2 ± 0.7	22.6 ± 0.6
$[C_{10}U_3]_3C_{10}$	49	$C_{10}U_3$	37.0 ± 0.6	24.2 ± 0.3

Table 1. Library of oligonucleotides sequences and data showing melting (T_m) and annealing (T_a) temperatures for the sequence $[C_nU_3]_3C_n$ including the average for the melting and annealing temperatures and standard deviation for 3 cycles.

*nd: not discernible. No i-motif formation was measured for this sequence.

2.2.3 Conclusions

Dahma's group have studied the effects of the stability of i-motif structure by characterizing the sugar interactions, and they have observed a weak or not distinguishable UV melting temperature for the RNA suggesting a lower stability (90), and Leroy *et. al.* have showed smaller CD signal suggesting lower thermal stability. However, our findings using UV spectroscopy at pH 5.5 have showed all the sequences from our library fold into i-motif structure but C_1U_3, and we have observed an increase in the T_m for all cytosine-rich sequences when the number of cytosines per tract increase, the T_a for all the sequences increased also, but after a number of six cytosines per tract length this increment is reduced. TDS results indicated a minimum of two cytosines per tract length at pH 5.5 in order to fold into i-motif structure. Comparing the DNA version of our sequence or C_nT_3 we found

using UV spectroscopy the T_m and T_a have shown higher thermal stability than C_nU_3, on the other hand using TDS we found for both C_nT_3 and C_nU_3 i-motif was the dominant secondary structure for all the sequences except sequences with one cytosine per tract length. The level of hysteresis for both C_nT_3 and C_nU_3 reached a higher point with sequences with 6 or more cytosines per tract length, but overall the degree of hysteresis was higher for the C_nU_3 suggesting less thermal stability than the C_nT_3.

2.3 Structure studies with Circular Dichroism at pH 5.5

2.3.1 Introduction

Having found a folded secondary structure with UV, we have further characterized our sequences from our library (Table1) using circular dichroism (CD) technique performed at sodium cacodylate buffer pH 4 to 7 at room temperature. By using this biophysical technique, we expected to know more about the conformational properties of the cytosine–rich sequences from our library. First, we have studied all the sequences (Table1) versus pH from 4 to 7 by using CD, but for the sequence $C_{10}U$ samples were prepared at pH from 4 to 6, due to resource constraints. Our CD experiments required a higher concentration of RNA (10 µM) than UV experiments (2.5 µM), however, we prepared samples at the most acidic pH since i-motif structure is more stable at this condition. Second, we have analysed CD spectra obtained at 288 nm since signal at this wavelength indicated the formation of i-motif structure which CD spectrum has a dominant and positive peak at 288 nm and a negative peak at 260 nm (105), also, the maxima ellipticity has been analysed reflecting the highest CD signal. We have observed how the pH affect the maxima ellipticity. Moreover, since CD spectra showed a shift up at 200 nm, data was analysed also. Finally, transitional pH (pH_T) was calculated from inflections points of fitting ellipticity at 200 nm, ellipticity at 288 nm and the maxima ellipticity (Table 2).

2.3.2 Results and discussion

According to the literature i-motif structures show a characteristic CD signal at 288 nm (106), then, the transitional pH was also calculated from inflection points by plotting data at 288 nm. At pH 5.5, all the sequences from the library showed a signal at 288 nm that was reflected by a positive band at this wavelength corresponding to i-motif structure except for the sequence C_1U_3 showing no discernible point (Appendix A3-a) and C_2U_3 did not show the peak at 288 nm, but at 281 nm, then signal shifted to the lower wavelength corresponding to the most basic pH (Figure 2.4 a), and C_4U_3 at pH 4 showed the folded structure at 288, also at pH 4.5 and pH 5.0 and unfolded at 5.5. (Appendix A3-d). Then, C_3U_3 showed the folded spectra until pH 4.5 then started to unfold at pH 5.5 corresponding to the preference of i-motif structure for acidic environments (Figure 2.5). The same for C_5U_3 and C_6U_3 showing the folded structure until pH 5.5 an unfolded at pH 6 shifting the signal to the lower wavelength (Figure 2.4 c and appendix A3-e, f). C_7U_3, C_8U_3 and C_9U_3 showed the folded structure until pH 5.5 and unfold at pH 6 (Appendix A3-h, i). Then, they are some difference in our CD results comparing with the CD of DNA equivalent sequence, at pH 5.5 for C_1T_3 we have the same results (Appendix A3-a), but our C_2U_3 did not fold into I-motif at 288 nm while the DNA sequence C_2T_3 did fold (Figure 2.4 a and b). While C_5T_3 for the DNA sequence has showed a peak at 288 nm at pH 7.0 (Figure 2.4 d) (13), our C_5U_3 has showed the peak at pH 5 and pH 5.5 (Figure 2.4 c**,** 2.5**)**. $C_{10}T_3$ was folded at pH 7.0 (Figure 2.4 d) (13) while $C_{10}U_3$ has folded at pH 5.5 and unfolded at 6.0, 6.5 and 7.0 (Figure 2.4 c and 2.5). These findings suggested sequences with more than 3 cytosines per tract length fold into i-motif and continue stable at pH 5.5. Comparing both C_nT_3 and C_nU_3, at pH 5.5 we found the CD signal shifted up at 200 nm for the C_nU_3 as the number of cytosine per tract increase for all the sequences. For example, C_5U_3 and $C_{10}U_3$ (Figure 2.4 c and e). But, the most remarkable difference was found at the basic pH 6.0, 6.5 and 7.0 for C_5U_3 and C_5T_3 (Figure 2.4 c and d), and $C_{10}U_3$ at pH 6.0 has a higher CD signal in compare with $C_{10}T_3$ at the same pH.

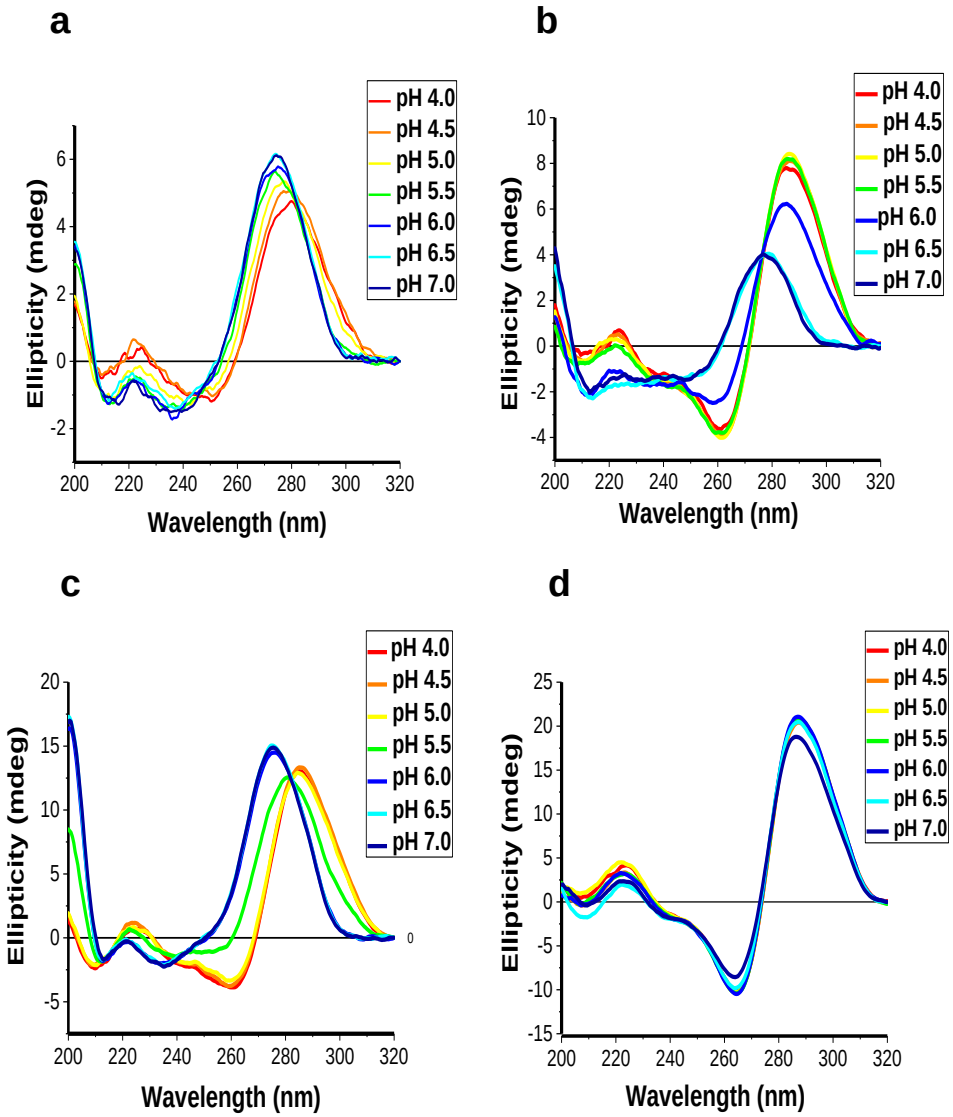


Figure 2.4 Circular Dichroism of 10 µM RNA or DNA in sodium cacodylate buffer (10 mM pH 5.5) and NaCl (100 mM) (**a**)C_2U_3(RNA), (**b**)C_2T_3(DNA), (**c**)C_5U_3(RNA), (**d**)C_5T_3(DNA)

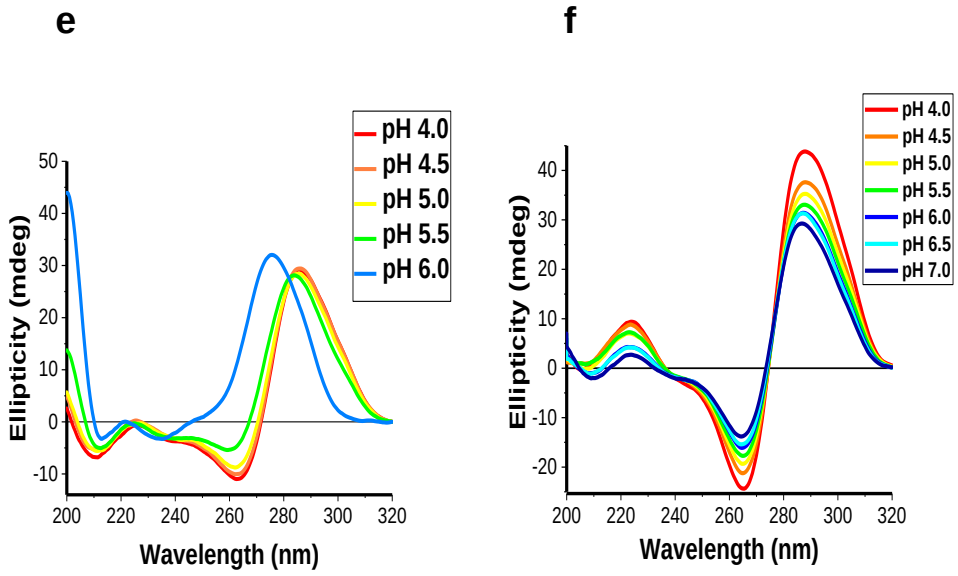


Figure 2.4 Circular Dichroism of 10 µM DNA or RNA in sodium cacodylate buffer (10 mM pH 5.5) and NaCl (100 mM) (**e**)$C_{10}U_3$(RNA), (**f**)$C_{10}T_3$(DNA).

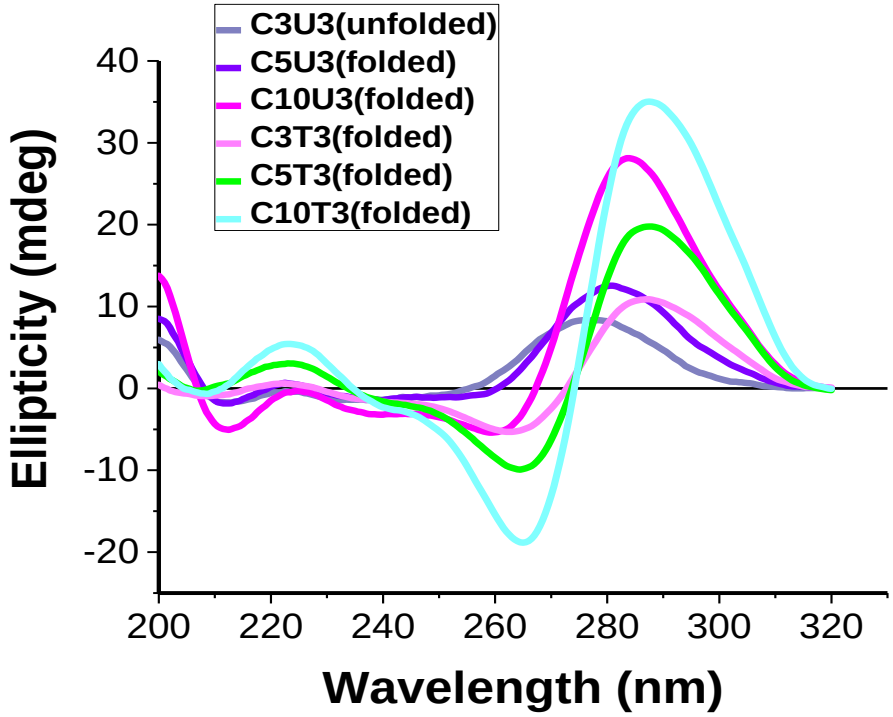


Figure 2.5 Circular Dichroism of 10 µM DNA or RNA in sodium cacodylate buffer (10 mM pH 5.5) and NaCl (100 mM) C_3U_3 (RNA), C_3T_3 (DNA), C_5U_3 (RNA), C_5T_3 (DNA), $C_{10}U_3$ (RNA), $C_{10}T_3$ (DNA).

We have calculated the pH_T transitional pH from the inflection points at 200 nm by plotting data and then fitted into a sigmoidal graphic since some of our sequences has showed a shift up in the CD signal at this wavelength (Figure 2.6) even though i-motif structure spectra show no peaks at this wavelength, our aim was to rule out or not the presence of other secondary structure that show signal at this amplitude. Some secondary structures have shown CD signal at this wavelength (105). All our cytosine rich sequences showed a signal at 200 nm, but C_1U_3 showed no discernible inflection points (Table 2 and Appendix A4-a). According to Allen and co-workers, the CD signal region around 200 nm and 230 nm seemed to be an effect of base compositions at least in DNA (129). Also, Well and collaborators have found bands at 210 nm and 240 nm in a study conducted in RNA and DNA double stranded suggesting these bands are related to the geometry of the base pairs of the oligos on their A and B forms (130). At 200 nm, the pH stability increased with the number of cytosine per tract length for example C_5U_3 showed a pH_T 5.5 that meant the secondary structure has changed from the folded state to the unfolded at pH 5.5 (Figure 2.6 and Table 2**)**. C_8U_3 and C_9U_3 showed pH_T 5.7 (Figure 2.6, Table 2 and appendix A4-h and i**)**. Since we could not have the CD signal of $C_{10}U_3$ at pH 6.5 and 7, and in order to predict the pH_T of $C_{10}U_3$ at 200 nm we plotted the data from all the sequences at 200 nm and fitted in a sigmoidal graphic to calculate the pH_T 5.6. It was appropriate to do this because the pH_T for all the sequences shown a gradual increment as the number of cytosines per tract length increase from 2 to 9 cytosines the pH_T was from 5.3 to 5.6. This finding suggested the pH_T for the $C_{10}U_3$ would be around 5.6 (Figure 2.6).

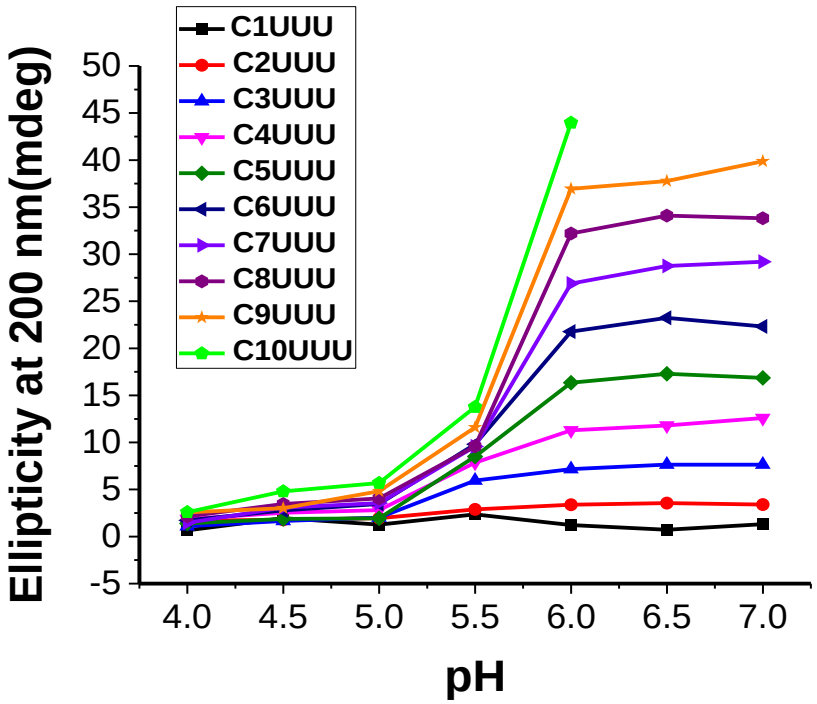


Figure 2.6 Ellipticity at 200 nm versus pH of all our sequences from our library at 10 µM RNA in sodium cacodylate buffer (10 mM pH 5.5) and NaCl (100 mM).

Even though $C_{10}U_3$ was performed only from pH 4 to pH 6, we observed the folded structure at pH 5.5 and unfolded at pH 6.0 (Appendix A5-e). $C_{10}U_3$, C_9U_3 and C_8U_3 data at 288 nm was fitted to a sigmoidal graphic to predict the pH_T of 5.6 for $C_{10}U_3$ (Appendix A5-j and Figure 2.7). As the cytosine tract length increase to seven cytosines the folded structure was present until pH 5.5 then, unfolded at pH 6. The stability of the folded structure increase as the number of cytosines per tract length until sequences with tract lengths of five showed a pH_T of 5.5 (Table2). This finding might suggest there is no change in the pH_T after sequences with five cytosines or more per tract. We have concluded that the increase in the number of cytosines per tract does not show a tendency towards pH 7.0, therefore not any of the sequences from our library will be stable at neutral pH.

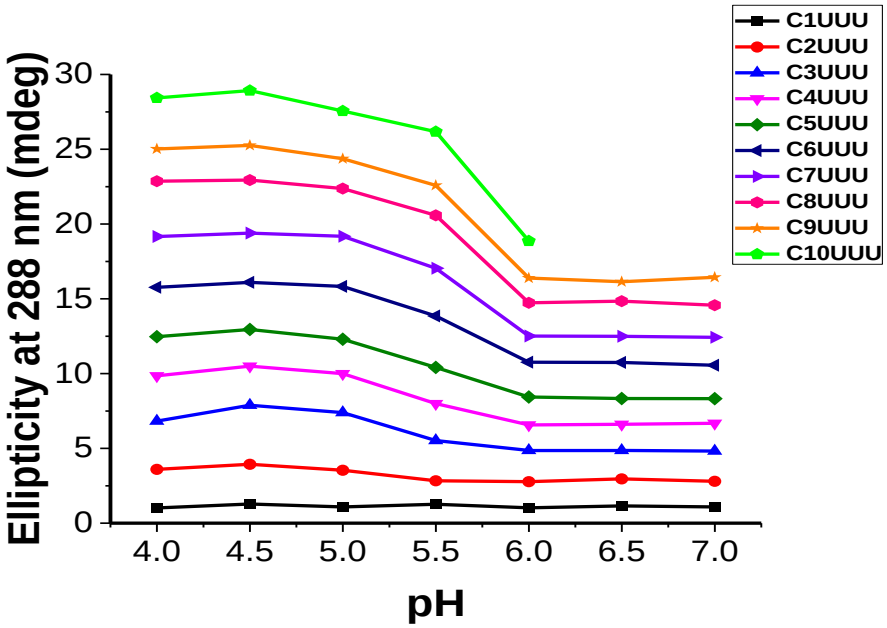


Figure 2.7 Ellipticity at **288** nm versus pH of all our sequences from our library at 10 µM RNA in sodium cacodylate buffer (10 mM pH 5.5) and NaCl (100 mM).

The transitional pH was calculated from the inflections points of fitting the maxima ellipticity in a sigmoidal graphic **(**Figure 2.8). All our sequences showed a CD signal, but not the sequence C_1U_3 (Table 2 and Appendix A6-**a**). The stability at the maxima ellipticity increased when the number of cytosines per tract length increased. For example, sequences with five cytosines the secondary structure changed from folded to unfold at pH_T of 5.7 (Appendix A5-e). $C_{10}U_3$ showed a pH_T of 5.6 when its data was plotted with C_9U_3 data in a sigmoidal graphic to predict a more accurate calculation (Appendix A5-j and Table 2). All sequences with more than five cytosines showed a minimal increase of pH_T around 5.6 suggesting the pH_T reached a plateau and the stability was not increased (Table2, Figure 2.8). These findings might suggest sequences with 5 or more cytosines per tract length fold into i-motif and maintain stable at pH 5.5.

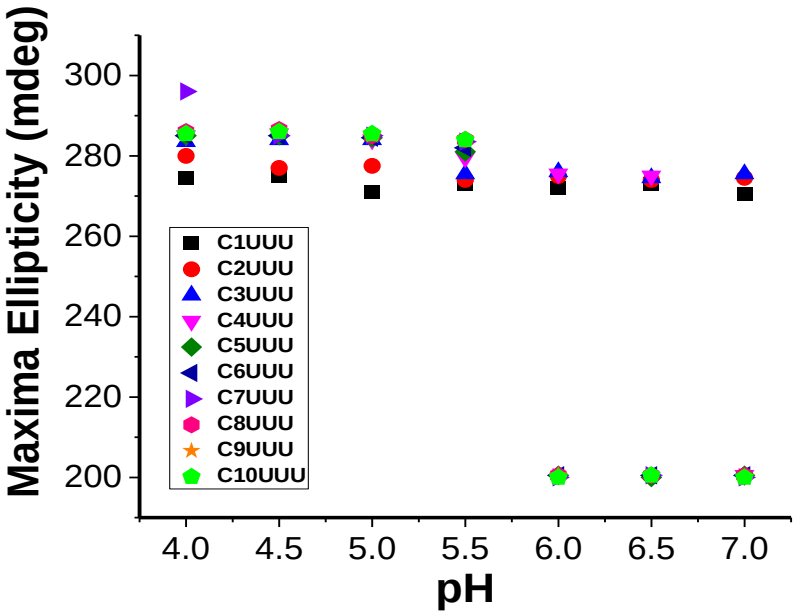


Figure 2.8 Maxima Ellipticity versus pH of all our sequences from our library at 10 μM RNA in sodium cacodylate buffer (10 mM pH 5.5) and NaCl (100 mM).

Sequence 5′- 3′	Bases	Notation	pH_T			
			Maxima ellipticity	Ellipticity 288 nm	Ellipticity 200 nm	Average/st. deviation
$[C_1U_3]_3C_1$	13	C_1U_3	nd	nd	nd	nd
$[C_2U_3]_3C_2$	17	C_2U_3	5.3	5.0	5.3	5.2 ± 0.2
$[C_3U_3]_3C_3$	21	C_3U_3	5.4	5.5	5.3	5.4 ± 0.2
$[C_4U_3]_3C_4$	25	C_4U_3	5.4	5.4	5.4	5.4 ± 0
$[C_5U_3]_3C_5$	29	C_5U_3	5.7	5.5	5.5	5.6 ± 0
$[C_6U_3]_3C_6$	33	C_6U_3	5.6	5.6	5.6	5.4 ± 0.3
$[C_7U_3]_3C_7$	37	C_7U_3	5.6	5.6	5.6	5.6 ± 0.05
$[C_8U_3]_3C_8$	41	C_8U_3	5.6	5.6	5.7	5.6 ± 0.05
$[C_9U_3]_3C_9$	45	C_9U_3	5.7	5.6	5.7	5.6 ± 0.05
$[C_{10}U_3]_3C_{10}$	49	$C_{10}U_3$	5.6	5.6	5.6	5.6

Table 2. The pH_T of the ellipticity at 200 nm, ellipticity at 288 nm and maxima ellipticity for all c-rich sequences from our library at pH 5.5

*nd: Not discernible inflection point.

2.3.3 Conclusions

Our findings from the CD spectroscopy have suggested i-motif structure was able to form with a minimum requirement of three cytosines per tract since all the sequences with three cytosines or more were able to fold into i-motif structure at 288 nm. TDS at pH 5.5 results shown sequences with two cytosines to ten cytosines for tract length were able to form i-motif structure. However, we might keep in mind UV melting/ annealing profiles and TDS spectra were performed at lower temperatures 4°C that tend to stabilise the structure (16), and CD technique was at room temperature which destabilise the structure. We also observed how the pH affect the maxima ellipticity for all our sequences. At the most basic pH the maxima ellipticity showed around 270 nm which mean the structure is unfolded, and as the pH decreased to the acidic the maxima ellipticity shifted toward 288 nm indicated the structure is folding into i-motif. From the literature, while the results from CD experiments with DNA sequence have concluded at pH 5.5 all the sequences were able to fold at 288 nm except the sequence with one cytosine, and the minimum of two cytosines per tract length to form i-motif (13), our results with CD experiments with RNA have observed at pH 5.5 a minimum requirement of three cytosines to fold into i-motif at 288 nm at 20°C. While the number of cytosines per tract length increase the T_m increase (Table1), the transitional pH_T reached a plateau in sequences with 5 or more cytosines per tract length (Table2). This finding might suggest that sequences with five or more cytosines are able to fold into i-motif and maintain stable at pH 5.5.

2.4 Thermal profile studies with UV spectroscopy at pH 7.4

2.4.1 Introduction

After our results from UV melting/annealing profile and TDS at pH 5.5 have suggested the minimum requirement of two cytosines per tract length to observe the folded and unfolded structure, we further characterized our cytosine-rich sequences at different pHs from 4 to 7 using CD at room temperature and found a minimum requirement of 3 cytosines per tract length for the structure to fold and unfold at the most acidic pHs knowing these pH conditions were optimal for the formation of i-motif structure, then, we have analyzed the sequences from our library to observe whether any of the sequences would able to fold into i-motif at physiological pH, but at lower temperatures (Table1).

2.4.2 Results and Discussion

At pH 7.4, UV melting/annealing profile have not shown any transition or melting point discernible for all the sequences from our library (Appendix A7 a-j, Figure 2.9). According to the literature DNA i-motif structure is prone to form in the presence of 5 or more cytosines at pH 7.4 (13) but we have concluded even increasing the length of the cytosines per tract, we could not observe a stable RNA i-motif structure at physiological pH on any of the sequences form the library.

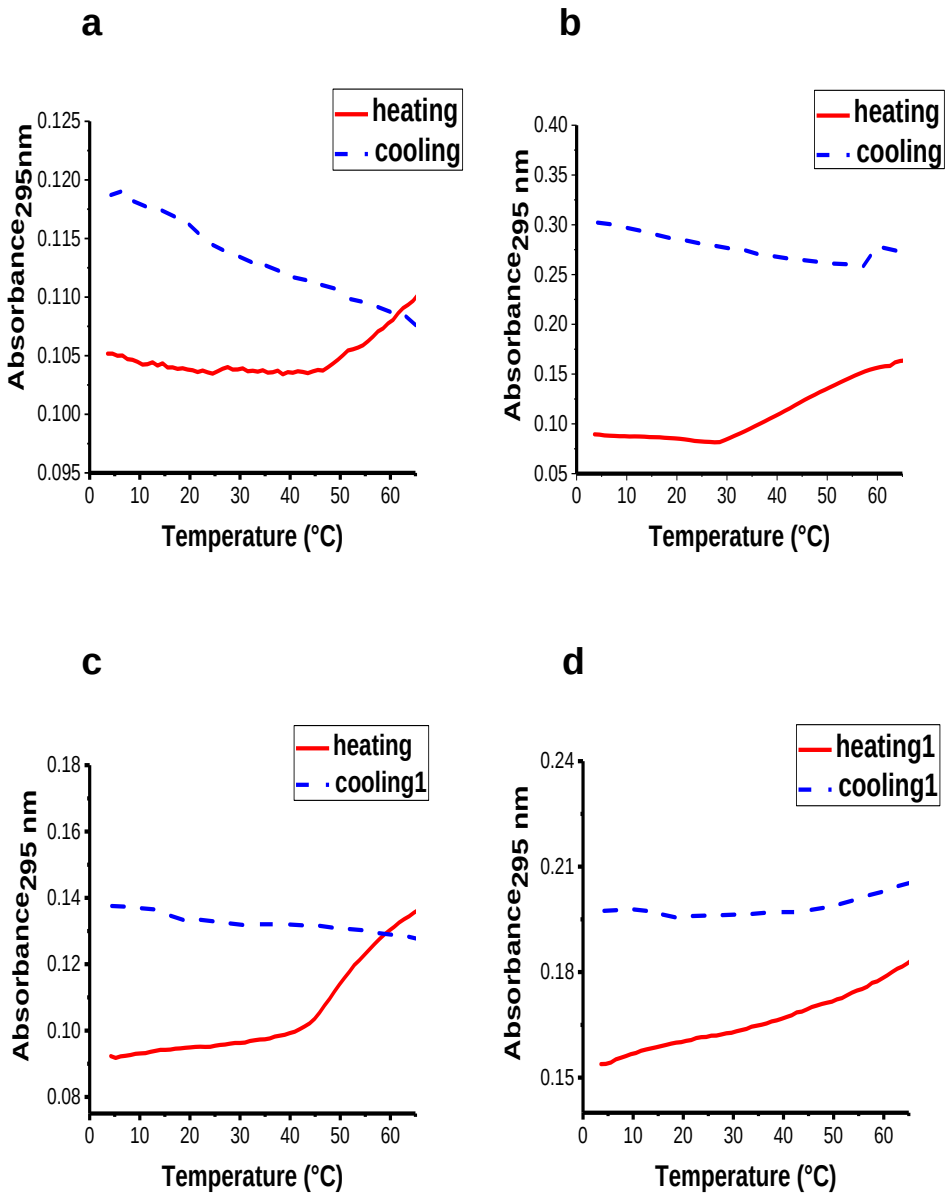


Figure 2.9 UV melting/annealing representative profiles of 2.5µM RNA in sodium cacodylate buffer (10 mM pH 7.4) and NaCl (100 mM) (**a**)C_1U_3, (**b**)C_2U_3, (**c**)C_5U_3, (**d**)$C_{10}U_3$.

2.4.3 Conclusions

From the literature DNA sequences were able to fold into i-motif structure at physiological pH without using any molecular crowder with a minimum requirement of five cytosines per tract length, and the T_m and T_a increased when the number of cytosines increased per tract length in general (13). In contrast with our results, at pH 7.4 not any of the cytosine-rich sequences from our library fold into i-motif structure (Appendix A**7**). From our knowledge, there are not examples of RNA i-motif at physiological pH in the literature. Lacroix *et. al.* have found the DNA i-motif can form i-motif structure at acidic and even neutral pH, but they observed RNA i-motif was not able to form in cytosine rich sequences above pH 6 *in vitro*, suggesting the RNA i-motif structure doubtful to find *in vivo* at neutral pH (23). Collin *et. al.* have observed the substitution of one DNA residue in one position of the sequence d(TC_5) for an RNA decrease the $\boldsymbol{T_m}$ in 6.5°C reducing the thermal stability of i-motif (89). Wright *et. al.* have observed comparing their sequences at two different pH 5.5 and 7.4 that the degree of hysteresis versus cytosine tract length was lower at pH 5.5 than 7.4 suggesting the kinetic process was more complex and slower at pH 7.4 (13). However, recent studies conducted by Zeraati and co-workers have reported an antibody fragment (iMab) that recognizes i-motif structures allowing the identification of i-motifs in the nuclei of human cells showing the existence of RNA i-motif structure (22). Our finding for the C_nU_3 at pH 5.5 have revealed a higher degree of hysteresis in compare with C_nT_3 that might suggest at pH 7.4 the process will be more complex for all our sequences. Also, we have observed from our CD results at pH 5.5 that does not matter how many cytosines have per tract the sequences from the library have, it will never be a sequence stable at neutral pH. We proposed to use a molecular crowder to mimic cell conditions that tend to stabilize i-motif (122) which potentially can stabilize RNA i-motif at physiological conditions.

Chapter 3: Investigating and characterizing RNA i-Motif in the sequence AT5G08230.1

3.1 Introduction

We have replicated and characterized the sequence $[C_nU_3]_3C_n$ at two different pH 5.5 and 7.4, and we have found that none of the cytosine-rich sequences from our library (Table1) can form i-motif at physiological pH. However, at pH 5.5, using a UV spectroscopy we have observed an increase of the T_m when the number of cytosines per tract increases. The degree of hysteresis for the melting/annealing curves also increases with the number of cytosines per tract length. TDS results indicated a minimum requirement of two cytosines per tract length in order to fold into i-motif structure. The results using CD indicated at pH 5.5 a minimum requirement of three cytosines to fold into i-motif at 288 nm, but it is important to mention all the cytosine-rich sequences have almost the same transitional pH_T regardless of the number of cytosines per tract (Table2).

Following our characterization of RNA i-motif, and knowing there are not studies about RNA i-motif in plants in the literature. We have proposed to study RNA i-motif in a plant sequence. We have known plants under stress like drought and salinity respond in different ways in order to survive (131), by changing their intracellular to different conditions of the environment. Since plants have to cope with changes like pH, temperature, salinity and others, we proposed i-motif would change under different conditions of stress environment and potentially alter biological functions. Eventually we will identify its role in response to stress conditions. Bearing this in mind we have studied a sequence in plants that can alter the stability by changing some conditions of pH, temperature, and adding a molecular crowding PEG 8,000, in this case, at two different percentages: 40 % and 20%. We have also used KCl to mimic drought conditions at two different concentrations 100 mM and 800 mM, knowing that plants cells under conditions of drought or in high salinity soils avoid cellular dehydration by increasing its ionic concentrations (particularly potassium cations) dramatically to 700 mM (31) this high K^+ level can stabilize i-motif formation.

We have studied the RNA sequence **AT5G08230.1** found in *Arabidopsis thaliana* using the same biophysical techniques we used with the RNA version of the sequence $[C_nU_3]_3C_n$ Ultraviolet UV, Circular Dichroism (CD) and Thermal difference spectra (TDS).

The RNA sequence **AT5G08230.1**: 5' -CCC-CCU-CCU-CAA-CCC-CCU-UCU-UCC-CCC-CCA-CCC-CC- 3'. This sequence has 35 bases and 26 of them are cytosine, 2 are adenine and 6 uracil. In collaboration with Ding's lab and using bioinformatics techniques, this specific sequence with five cytosine tract length was found and after a comprehensive assessment of RNA i-motif forming sequences. This i- motif forming sequence with five cytosines per tract length, one of the most stables, of *Arabidopsis,* was associated to plant functions which would have biological interest (132). The **AT5G08230.1** is associated with some biological process like flower development, regulation of transcription and mRNA processing. Also, it is cellular component of plants located in the nucleus and plasmodesma (133) and it is known by other names HUA2 LIKE 1, HULK1 and SL5. This group of genes are in charge of regulate the function and expression of some essential genes specifically HUA2 like Flowering Locus C (FLC) central repressor of flowering time (134), and AGAMOUS involved in reproductive development (133). According to Jali and co- workers, plants lacking the HUA2 gene expressed a moderate reduced plant stature rate, flowering time and leaf initiation development (133). Post transcriptional regulation is involved in the process of gene expression during development, and genes like FLC AND AGAMOUS participate actively in this process because they encode RNA-binding proteins (133). Considerable amount of work has been done on transcriptional regulation on the response to different pH. But, there is little in the literature about the mechanism of post-transcriptional regulation. We investigated the stability and folding potential at pH: 5.8 which plants normally thrive on. Additionally, we have studied this plant sequence under drought conditions by increasing the concentration of K^+ and crowding conditions by using PEG 8,000.

3.2 Structure studies with Circular Dichroism (CD) characterizing the sequence AT5G08230.1 RNA.

3.2.1 Introduction

In order to confirm this sequence was able to form i-motif structure this sequence was examined at different pH from 4 to 8 including 5.8 using CD.

3.2.2 Results and discussion

We obtained different conformational changes in the structure of the samples in the graph (Figure 3.1) showing a shift from right to left to the lower wavelength for every one of our sample with the different pH from 4 to 8. There is a correlation between the value of the pH of the samples and the graphics. The conformational changes increase as the value of the pH of the samples, and the CD spectra of the lowest values of the pH samples corresponding to 4, 4.5, and 5 showed a dominant positive peak around the 290 nm and a negative peak around 255 nm indicating the formation of cytosine-cytosine+ base pair confirming the preference of i-motif structure for acidic conditions (106) (Figure 3.1).

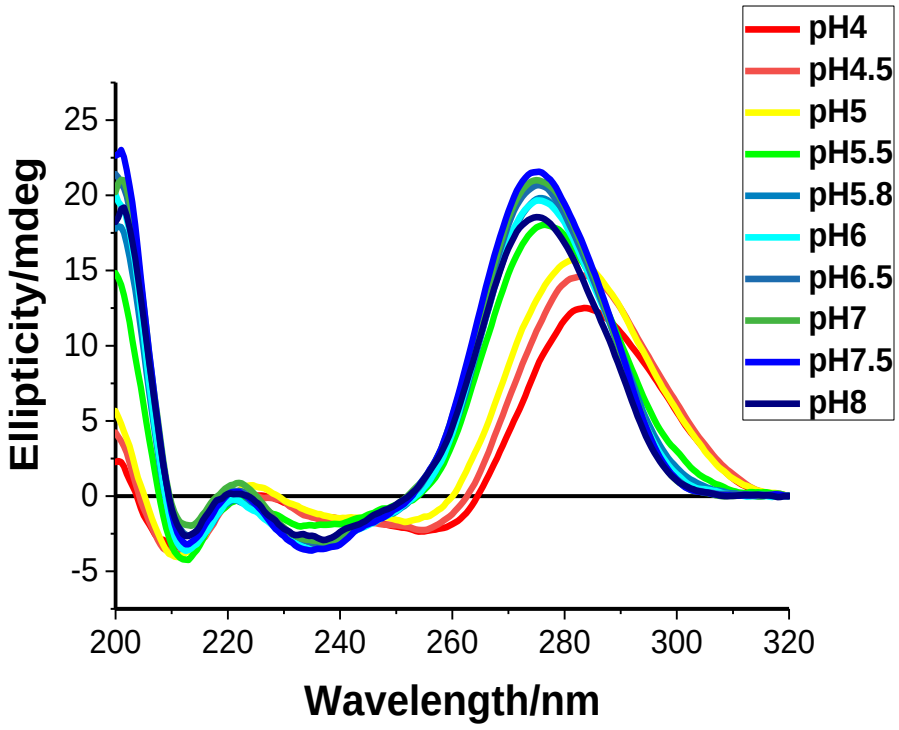


Figure 3.1. Circular Dichroism of 10 µM AT5G in sodium cacodylate buffer (10 mM) and KCl (100 mM) at different pH from 4 to 8.

The transitional pH (pH_T) for the maxima ellipticity, ellipticity at 288 nm, ellipticity at 275 nm and 200 nm were calculated by plotting those values versus pH in a sigmoidal curve, and fitted to calculate them. The transitional pH (pH_T) for the maxima ellipticity was 5.2 (Figure 3.2 a) (Table 3). The transitional pH (pH_T) for the ellipticity at 288 nm was 5.4 (Figure 3.2 b) (Table3). The transitional pH (pH_T) for

the ellipticity at 275 nm was 5.1 (Figure 3.2 c), and for the ellipticity at 200 nm was 5.3 (Figure 3.2 d) (Table 3).

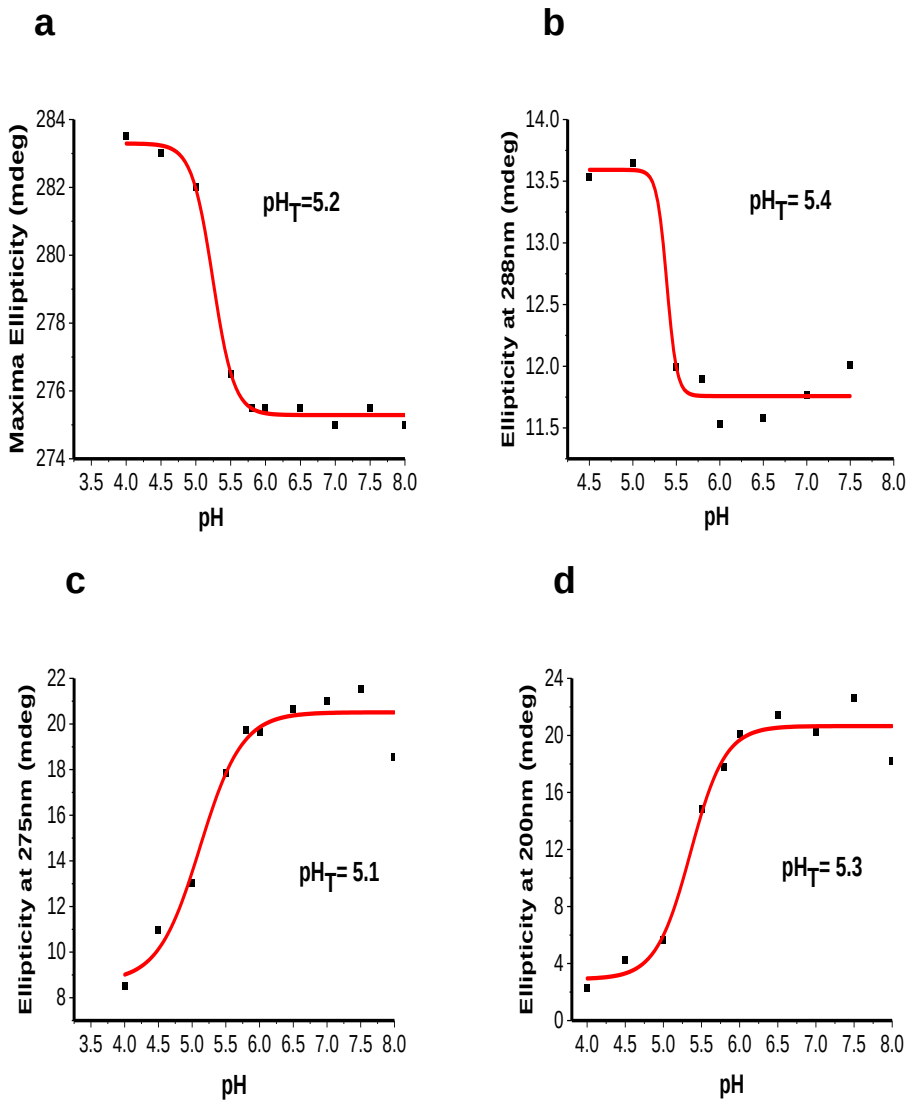


Figure 3.2 Circular Dichroism of 10 µM AT5G in sodium cacodylate buffer (10 mM) and KCl (100 mM) of (**a**) Max. Ellipticity vs. pH (b) Ellipticity at 288 nm vs. pH (c) Ellipticity at 275 nm vs. pH (d) Ellipticity at 200 nm vs. pH.

pH	λmax	Ellipticity at 275 nm	Ellipticity at 288 nm	Ellipticity at 200 nm
4	283.5	8.4	11.6	2.2
4.5	283	10.9	13.5	4.2
5	282	13.0	13.6	5.6
5.5	276.5	17.8	11.9	14.8
5.8	275.5	19.7	11.8	17.7
6	275.5	19.6	11.5	20.0
6.5	275.5	20.6	11.5	21.4
7	275	21	11.7	20.2
7.5	275.5	21.5	12	22.6
8	275	18.5	10.3	18.2

Table 3. Showing the effect of changing values of pH at Ellipticity at 200, 275 and 288 nm.

3.2.3 Conclusions

Our CD measurements with **AT5G08230.1 RNA** have indicated samples with pH 4.5, 5 and 5.5 folded into i-motif at 288 nm having the sample with pH 5 the highest CD signal. (105) (Figure 3.1), then from pH 5.5 and 5.8 showed a decreased in the CD signal. They also showed a positive peak around 275 nm and negative peak around 215 nm which would correspond to an unfolded conformation (75, 105) or another structure but not i-motif. All the transitional pH (pH_T) or the midpoint where the transition is between the unfolded and folded stage for the maxima ellipticity, ellipticity at 288 nm, 275 nm and 200 nm were around 5.2 confirming the preference of i-motif structure for acidic conditions. When values for the maxima ellipticity was plotted versus pH, we have observed samples with the most basic pH showed a value around 275 nm which mean the structure is not folded (Figure 3.2 a). On the other hand, when the values for the ellipticity at 288 nm were plotted versus pH, samples with the most acidic pH 4 and 4.5 showed the highest values (Figure 3.2 b). Values for the ellipticity at 275 nm were plotted versus pH, the CD signal increased as the pH values increased

reaching a plateau at pH 6.5 after what the increased was minimal (Figure 3.2 c) and (Table3). The values for the ellipticity at 200 nm were plotted versus the pH, samples with the highest pH showed the highest CD signal (Table3) and (Figure 3.2 d).

3.3 Structure studies with Circular Dichroism- titration with K^+

3.3.1 Introduction

CD assay titration with K^+ have allowed us to see conformational isomerization of the oligo after adding small amounts of KCl that induced conformational isomerization in the molecule of RNA. The CD spectral changes make it possible to spot gradual changes in the oligo (105). According to the literature, cations have an effect on i-motif forming DNA sequences, for example Famulok and co-workers have observed that two guanine or cytosine rich sequences can form a Watson and Crick double helix or individually form into a G-quadruplex and i-motif structure in the presence of H^+ and K^+ (75). In fact, they have reported their DNA cytosine-rich sequence fold into i-motif at pH 4.5 in the presence or not of KCl (75). Additionally, Waller and co-workers have observed the use of silver cations in a cytosine-rich sequence to form i-motif structure at biological pH. CD titration with Ag^+ was conducted to observe the effect of this cation on conformation of the molecule of DNA. They have observed a gradual shift of the CD signal in the positive peak at 275 nm after gradual addition of Ag^+ to the higher wavelength at 288 nm which indicated the presence of i-motif structure (66). On the other hand, plants have adapted in the course of evolution to a wider range of mechanism in order to survive. In fact, there is clear evidence of the role that mineral nutrients play in plant stress resistance (31). Plants under regular conditions of the environment have between 100 and 200 mM K^+ in the cytosol under stress conditions like drought or salinity plants try to avoid cellular dehydration by increasing their amount of K^+ in the cells up to 500 mM or more (135). Considering K^+ the most abundant cation in plants and the important role that play in the survival of plants under abiotic stress, and keeping in mind we have already characterized this plant sequence using CD, as well as in the literature about i-motif and conformation of RNA.

3.3.2 Results and Discussion

CD spectra was recorded from the samples, first without any KCl, then small amounts (1µL) of KCl (3M) were added gradually. Every 1µL was an equivalent of 17 mM of KCl so, 2µL =34 mM and so on. The maximum amount of KCl added could not exceed a volume that could dilute more than the 10% of the total volume of the buffer 175 µL. Therefore, the last addition was 17µL (306 mM). CD spectra of the sample before any addition started as an unfolded random coil with a major and positive peak at 278.5 nm and a negative and weak peak around 235 nm suggesting the presence of a folded structure but not i-motif structure (Figure 3.3)

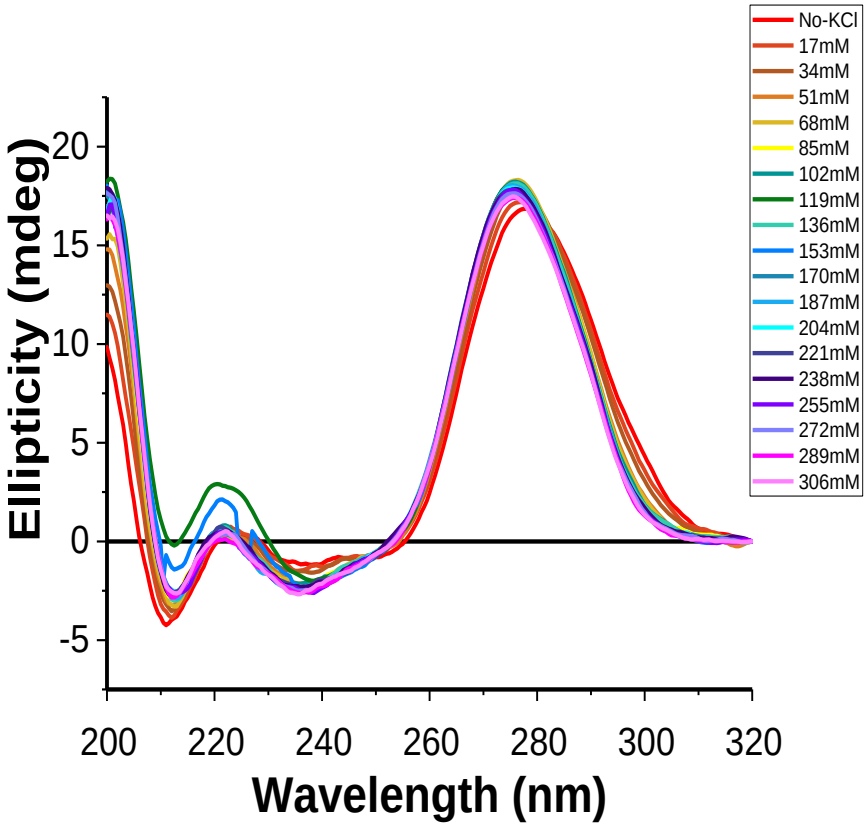


Figure 3.3 Circular Dichroism-titration 10 µM AT5G with KCl (3M) in sodium cacodylate buffer (10 mM a pH 5.8).

KCl/mM	ƛmax/mdeg	Ellipticity at 200 nm
No KCl	278.5	9.8
17	277.5	11.5
34	276.5	12.9
51	276.5	14.8
68	276.5	15.3
85	275.5	16.9
102	276	16.9
136	276	17.8
153	200	18.0
170	275.5	17.5
187	275.5	16.8
204	275.5	16.6
221	276.2	16.3
238	200.5	17.9
255	275.5	16.5
272	200	17.6
289	276	16.3
306	275.5	16.5

Table 4. Values for the Maxima ellipticity and Ellipticity at 200 nm obtained from the Circular Dichroism- titration of 10 µM AT5G with KCl (3M) in sodium cacodylate buffer (10 mM at pH 5.8).

Also, the CD signal with the highest value for maxima ellipticity 278 nm was taken from the sample without any addition of KCl. Then, in addition of up to 136 mM, the CD signal shifted from right to the left to the lower wavelength to 276 nm until the last addition of 17µL (306 mM) showed a decrease of the intensity of the signal with dominant positive peak is around 275nm (Table 4). We have observed small conformational changes, decrease in the signal until 170 mM where no further changes in the ellipticity were indicated, and ellipticity signal was around the same range of wavelength from 278 to 275 nm (Figure 3.4 a). We have observed the ellipticity at 200 nm presented some shifts up in the CD spectra. CD signal shifted up at 200 nm until 153 mM of KCl was added, then, CD signal has increased showing small changes suggesting reached a plateau (Figure 3.4 b) and (Table4)

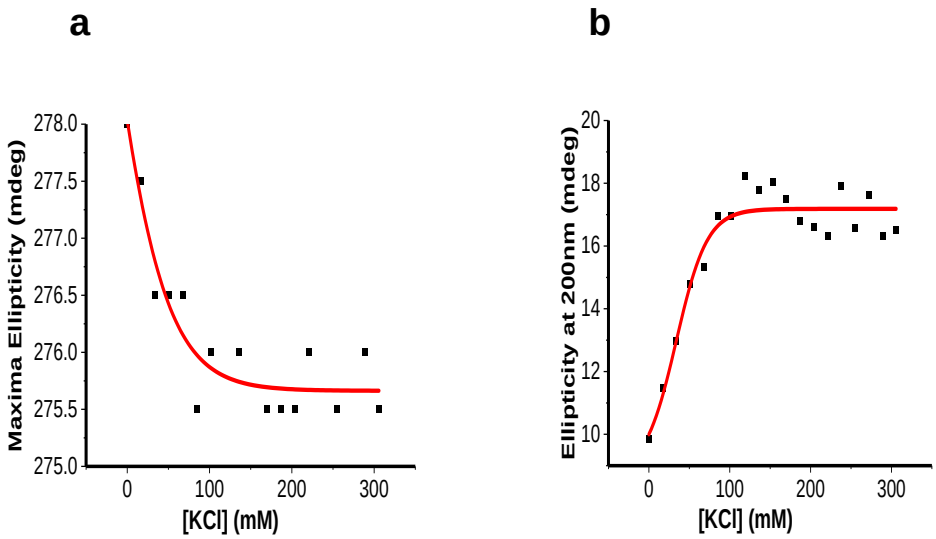


Figure 3.4 Circular Dichroism spectra of 10 µM AT5G (**a**) Maxima Ellipticity vs. KCl. (**b**) Ellipticity at 200 nm versus KCl.

3.3.3 Conclusions

CD spectra of **AT5G** at pH 5.8 started as an unfolded random coil with a dominant and positive peak at 278.5 nm and a second and negative peak around 235 nm indicating the presence of a folded structure but not i-motif structure (Figure 3.3). When the values for the maxima ellipticity were plotted versus KCl, and fitted in a sigmoidal curve, we have found small conformational changes indicated by a decrease in the signal until 170 mM of KCl was added. After that, the CD signal had changed, indicated by small conformational changes at the same range of wavelength from 278 to 275 nm (Figure 3.4 a). Values for the ellipticity at 200 nm were plotted versus KCl, the CD signal changed increasing as the amount of KCl increase indicated for small conformational changes (Figure 3.4 b) and (Table4).

3.4 Thermal profile studies with UV spectroscopy at pH 5.8

3.4.1 Introduction

From CD experiments with pH from 4 to 8, we have observed **AT5G08230.1 RNA** was able to fold into i-motif at the most acidic pH. UV spectroscopy was recorded to characterize and analyze the thermal profile of this sequence. We have used sodium cacodylate buffer at pH 5.8 because that is the pH of the soil where plants thrive normally (136, 137), and 100 and 800 mM KCl was used to check the influx of potassium at physiological concentration. By using the UV spectroscopy technique, we were looking to identify the melting and annealing profiles of the secondary structure present in the sample, and to observe how KCl at 100 mM affects the stability of RNA i-motif.

3.4.2 Results and Discussion

We have found under these conditions the sample showed two states: folded and unfolded for the melting and annealing curves (Fig. 3.5 a). And, the T_m melting temperature between the folded and unfolded stage of the oligo was the 12.1°C. The average for the 3 melting cycles was 12.1°C ± 1 and the T_a for the three cooling cycles is 12.4°C ± 0.5 (Table5). The thermal difference spectra (TDS) was obtained by registered the absorbance of the unfolded and folded states of the sample above and below their melting temperature. In this case the TDS thermal difference spectra at pH 5.8 at 20°C has shown a random coil shape while at 4°C is within the limits of the spectrum of a classical i-motif structure with one major peak at 240 nm, and a second negative peak at 290 (Fig. 3.5 b). UV spectroscopy at 800 mM KCl showed the average for the 3 melting curves at 5.7 °C ± 0.5, and the average for the annealing curves of 5.7°C ± 1.5 (Fig. 3.6 a). The TDS has showed at 4°C is within the limits of i-motif structure (105) (Fig.3.6 b).

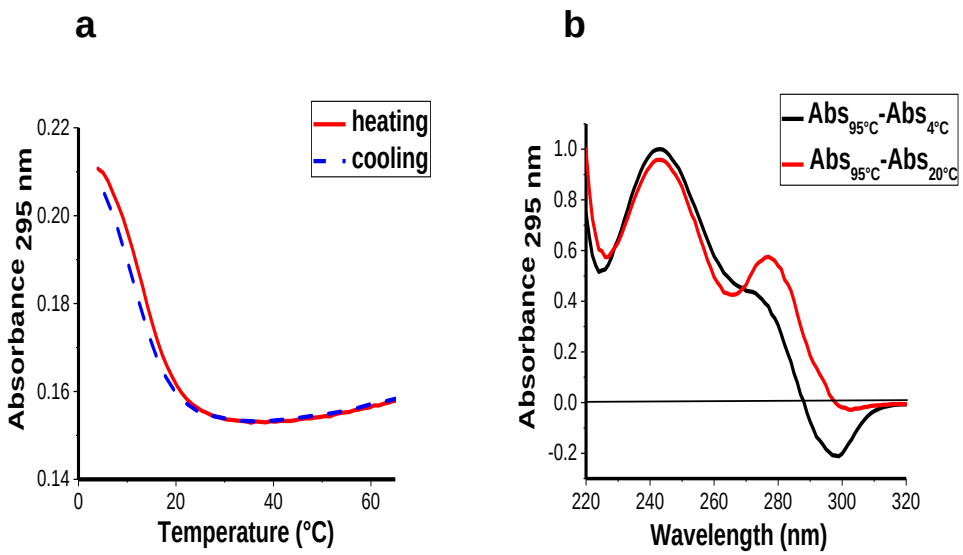


Figure 3.5 UV melting and annealing profiles of 2.5 µM AT5G RNA in KCl (100 mM) and sodium cacodylate buffer (10 mM pH 5.8) (**a**) Profile showing all heating and cooling cycles (**b**) The thermal difference spectra (TDS) calculated between 95 and 4°C

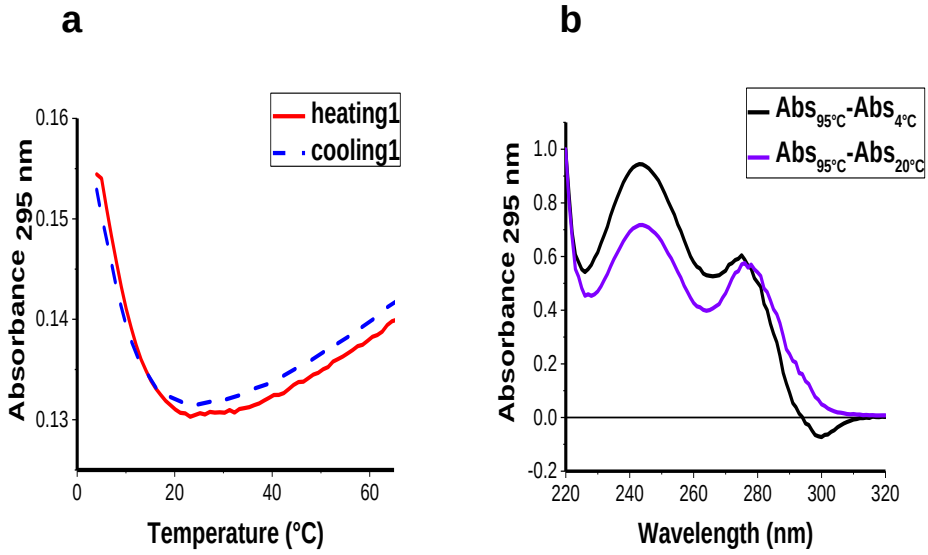


Figure 3.6. UV melting and annealing profiles (a), (b) The thermal difference spectra (TDS) calculated between 95 and 4°C.of 2.5 µM AT5G RNA in KCl (800 mM) and sodium cacodylate buffer (10 mM pH 5.8)

AT5G008230.1	pH 5.8	
	T_m /°C	T_a /°C
100 mM KCl	12.1 ± 1	12.4 ± 0.5
800 mM KCl	5.7 ± 0.5	5.7 ± 1.5

Table 5. UV- melting and annealing representative profiles showing the average melting temperature T_m *and* T_a for the 2.5 µM AT5G008230.1 (10 mM sodium cacodylate pH 5.8) and (100 and 800 mM KCl).

3.4.3 Conclusions

Comparing these results of the melting and annealing temperatures at pH 5.8 with our results from library of sequences at pH 5.5 (Table1), we have observed at 100 mM KCl the T_m and T_a have occurred at lower temperature 12°C suggesting a low thermal stability than the sequence C_5U_3 with a T_m of 21 ± 0.9 and T_a 19.4 ± 0.6. Our plant sequence AT5G has tract lengths of 5-5-7-5 cytosines and short loop lengths of 2 and 1 adenine and uracil intercalated inside the sequence. That would provide a thermal stability to i-motif structure according to the concept of shorter loops provide higher thermal stability to i-motif structure (14). The results for the annealing and melting profiles at 800 mM KCl were unexpected, and close to 4°C that is the temperature where the heating cycles started measuring the absorbance at 295 nm suggesting no transition (Fig. 3.6 a).

3.5 Thermal profile studies with UV spectroscopy mimicking crowding conditions using PEG 8,000 (40%)

3.5.1 Introduction

According to the literature, plant cells have 30-40% of molecular crowding under normal conditions (119). In order to mimic crowding conditions in plants cells PEG 8,000 (40%) was used as a molecular crowder. UV spectroscopy was used to determine both the melting(T_m) and annealing (T_a) profile for this sample with PEG which stabilize i-motif structure by increasing the melting temperature (64, 120).

3.5.2 Results and Discussion

At pH 5.8, both samples showed two state (folded and unfolded) melting and annealing curves. We already have the results for earlier experiments for the samples at 100 mM KCl the average melting temperature (T_m) was 12.1°C ± 1 and the annealing temperature (T_a) was 12.4 ± 0.5 (Figure 3.2.4 a) while the melting temperature (T_m) for the sample with PEG 8,000(40%) was 30.6°C ± 0.5, and the annealing temperature (T_a) was 38.3°C ± 7 (Fig.3.7a). Our findings have suggested the molecular crowder PEG stabilized i-motif structure in its folded state This condition made the average for melting/annealing temperatures higher (Fig.3.7 a) (Table 6). The thermal difference spectra (TDS) was taken by recording the UV-absorbance of the folded and unfolded state of the sample for above and below the melting temperature. TDS revealed a “classic i-motif structure” with a dominant positive peak at 240 nm and a negative peak at 297 nm. (Fig.3.7 b).

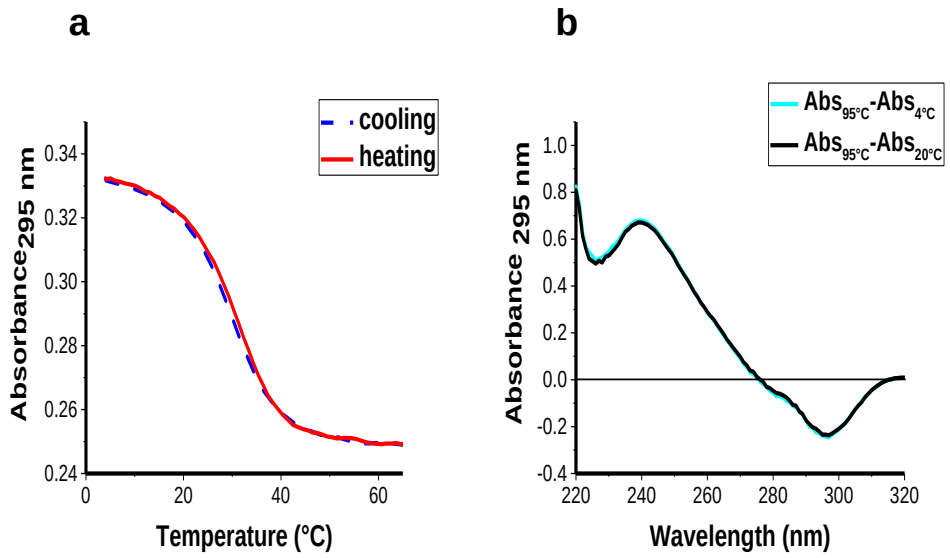


Figure 3.7 UV melting and annealing representative profiles of 2.5 µM AT5G RNA in KCl (100 mM), sodium cacodylate buffer (10 mM pH 5.8) and PEG 8,000(40%) (**a**) Profile showing raw data (**b**) The thermal difference spectra (TDS) calculated between 95 and 4°C.

AT5G008230.1	**pH 5.8**	
	***Tm*/°C**	***Ta*/°C**
No-PEG	12.1 ± 1	12.4 ± 0.5
PEG 8,000 (40%)	30.6 ± 0.5	38.3 ± 7

Table 6. UV-melting and annealing profile of AT5G in KCl (10 mM), sodium cacodylate buffer (10 mM pH 5.8) showing the different T_m and T_a average for two different samples under the same RNA conditions in the presence and absence of PEG 8,000 (40%).

3.5.3 Conclusions

Thus, the difference of T_m between the No-PEG and PEG 8,000(40%) solutions was 18.5°C, and 25.9°C for the T_a values. We have observed an increase of 40% of the melting temperature of the samples after the addition of PEG 8,000. This finding suggests that the addition of PEG 8,000 not only induces the complete folding of i-motif RNA at the pH 5.8, but also stabilizes i-motif structure. Having established the presence of secondary structures by UV spectroscopy, TDS were determined to confirm the presence of i-motif structure with a dominant positive peak at 240 nm and a negative peak at 297 nm (105).

3.6 i-Motif structure studies with PEG 8,000 (40%) using CD spectroscopy

3.6.1 Introduction

CD spectroscopy was used to further characterize the secondary structure involved in the melting and annealing curves. To mimic the crowding conditions of the cell, we analyzed the sequences both in the presence and absence of PEG 8,000 (40%).

3.6.2 Results and Discussion

CD spectra recorded for the sample without PEG shows a dominant positive peak at around 276 nm and a negative peak around 238 nm (Fig.3.8). It would correspond to a B-form variant structure like a hairpin, according to Kypr (106). The second sample was prepared with the same RNA and sodium cacodylate buffer conditions, but PEG 8,000 (40%) was added to the solution. CD spectra shifted to the higher wavelength from 276 to 288 nm, and showed a second and negative peak around 255 nm revealing the equilibrium shifted to the “classic i-motif structure” (Fig. 3.8). The melting temperature $\boldsymbol{T_m}$ at 30°C has shown similar values to the $\boldsymbol{T_m}$ of most of the DNA i-motif structures from the literature, for example Mergny *et. al.* have observed a $\boldsymbol{T_m}$ of 38°C at pH 5.6 in a i-motif forming sequence (11) and Leroy *et. al.* have found a $\boldsymbol{T_m}$ of 29°C at pH 6.4 (17).

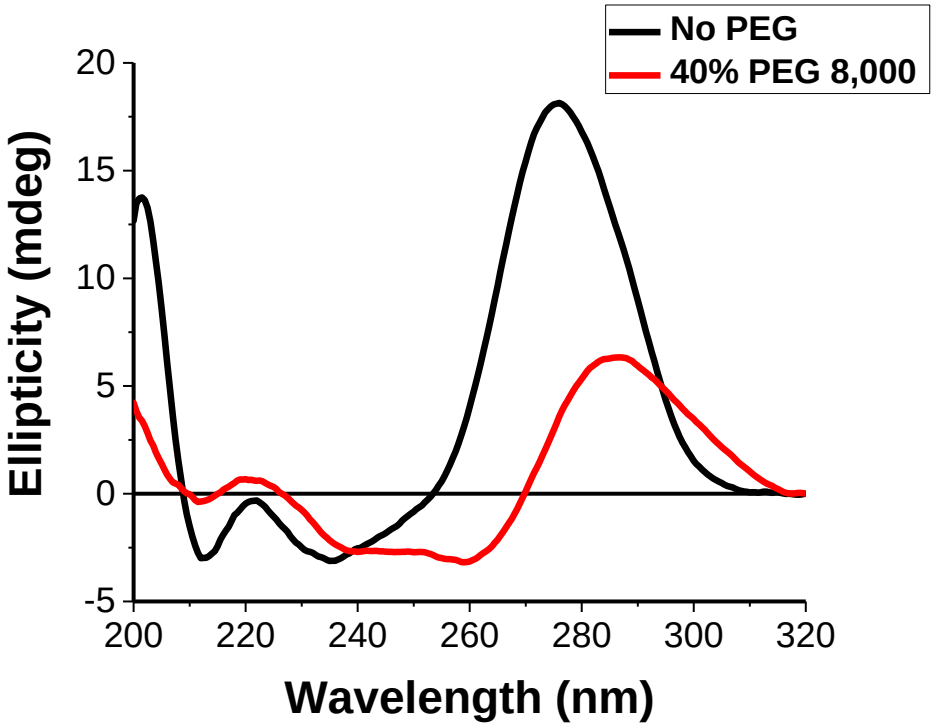


Figure 3.8. Circular Dichroism of 10 µM AT5G in the absence and present PEG 8,000 (40%) in sodium cacodylate (10 mM pH 5.8) with 100 mM KCl.

3.6.3 Conclusions

CD spectra of AT5G with PEG 8,000 (40%) confirm the presence of a classic i-motif structure with a dominant and positive peak at 288 nm and a second and negative peak around 255 nm. This finding has suggested that the addition of PEG 8,000 promotes the folding of i-motif RNA structure at pH 5.8, and stabilizes i-motif structure. It is important to mention our CD signals for the unfolded i-motif showed a higher ellipticity in compare with the folded i-motif structure, and our CD signal value is lower in compare with the CD signal for DNA i-motif of Kypr *et. al.* (105), but our conditions of buffer, pH and others are different with Kypr's conditions, so we might expect different results. But, our CD signal has shown some similarities with the i-motif structure of hTeloC with the same conditions from chapter 1 (Figure 1.14) obtained from Wright *et. al*. (13).

3.7 Control Experiments with pH

From the literature, we found that Arakawa and coworkers have observed cosolutes like PEG at higher percentage than 30% change the pH of the buffer solutions (138). Even though the data presented above has shown a small standard deviation, we were concerned that the PEG was changing the pH of the solutions, and had an effect in the melting temperatures. Bearing this in mind, we have performed a series of control experiments to check if the pH of the samples changed after they went through the heating cycles in the ultraviolet (UV) spectroscopy machine. We checked the value of the pH of the solutions after they were exposed to the heating cycles, and found the pH actually changed. We decided to perform some control experiments with pH. Our first set of control experiments was conducted with “dummy” samples without RNA, but buffer sodium cacodylate 10 mM, potassium chloride 100 mM and water. We measured the pH 5.87. Then, the PEG 8000(40%) was added, and pH was measured. The pH changed to 4.65 showing 1.22 units of difference. These results suggested the PEG was changing the pH of the sample without any heating cycle involved already. If the buffer at 10 mM was affected by the PEG, we decided to increase the concentration of the buffer to 50 mM, so at higher concentration we were expecting the changes would be minimal to none. A new PEG 8,000(60%) stock solution was prepared using the sodium cacodylate buffer at 50 mM. By using the buffer instead of water, what we were looking for was to be sure the pH of the final solution would not change. Then, we prepared a dummy sample without RNA with the 50 mM buffer, potassium chloride and water, the pH was measured: 5.83. We added at PEG 8,000(40%), and pH was measured: 5.93. There was a 0.1 units of difference. We would say there is not a significant difference, but it can make the difference for forming i-motif structure. Mergny and collaborators have observed how the change of pH can affect the thermal stability of i-motif, and found a decreased of minimum 20°C of the melting temperature T_m when the pH increased by 1 unit (11).

3.7.1 Thermal profile studies with UV spectroscopy with 100 mM KCl, PEG 8,000 (20%) and 800 mM KCl.

The first part of this experiment was conducted with sodium cacodylate buffer at 50 mM in the absence and presence of PEG to obtain a melting profile and uses to compare this data when we modified the conditions. The UV thermal profile has shown the melting T_m 16.8°C and annealing T_a 15.8°C for the sample without PEG (Fig.3.9 a). The thermal difference spectra TDS at 4°C showed within the limits of i-motif structure with a major positive peak around 244 nm, and second negative peak at 296 nm (Figure 3.9 b).

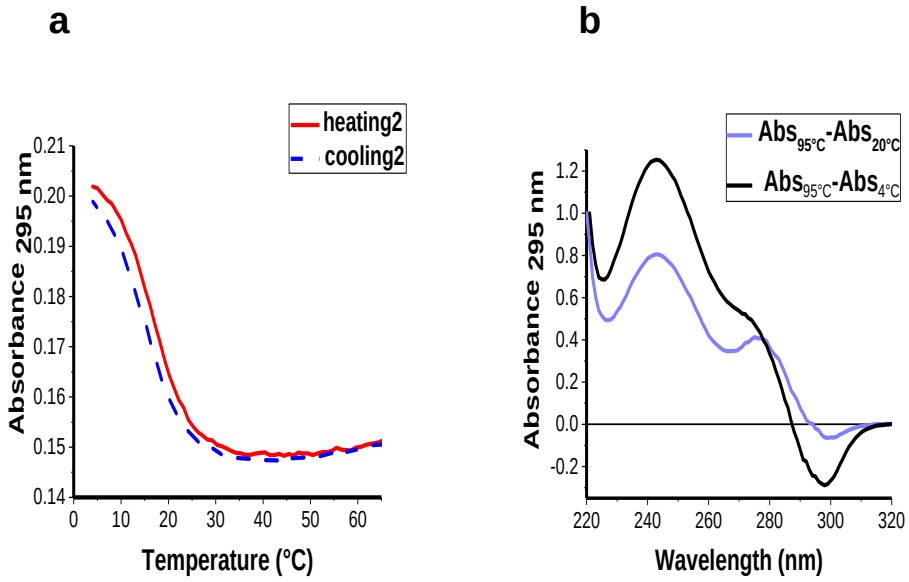


Figure 3.9. UV melting and annealing profiles (**a**) and the thermal difference spectra (TDS) (**b**) calculated between 95 and 4°C of 2.5 µM AT5G RNA in KCl (100 mM) and sodium cacodylate buffer (50 mM pH 5.8).

UV spectroscopy with PEG 8,000 (20%) has shown the melting (***Tm***) 24°C and annealing temperatures (T_a) 28°C (Fig.3.10 a). We have observed an increase of the melting and annealing temperatures suggesting the molecular crowder at 20% stabilized i-motif structure by increasing melting temperature in 7 °C in compare with the results of the AT5G without PEG. Also, the T_m increased after every cycle having 18.14 °C for cycle 1, 25.4°C and 28.3°C for cycle 2 and 3 respectively. The

TDS showed a classic i-motif structure with a major positive peak at 241 nm and a second peak at 296 nm (Fig.3.10 b).

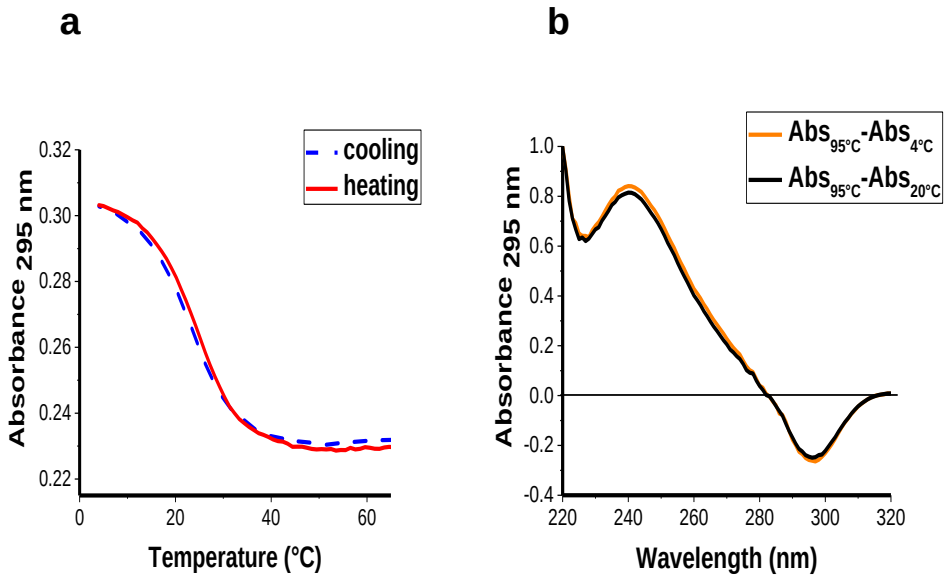


Figure 3.10 UV melting and annealing representatives profiles (**a**) and the thermal difference spectra (TDS) (**b**) calculated between 95 and 4°C of 2.5 µM AT5G RNA in KCl (100 mM) and sodium cacodylate buffer (50 mM pH 5.8) and PEG 8,000(20%).

Having characterized the UV thermal profile in the absence of PEG at 100 mM KCl (Figure 3.9 a**)**, we increased the concentration of potassium chloride at 800 mM to mimic drought condition in plants. UV melting profile for the AT5G with 800 mM KCl revealed a transition approaching the starting temperature of UV spectroscopy at 4°C for the melting/annealing curves (Fig.3.11a). The TDS spectra at 4°C showed a shape within the limits of i-motif structure with a major positive peak around 240 nm and a second peak around 290 nm (Fig.3.11 b).

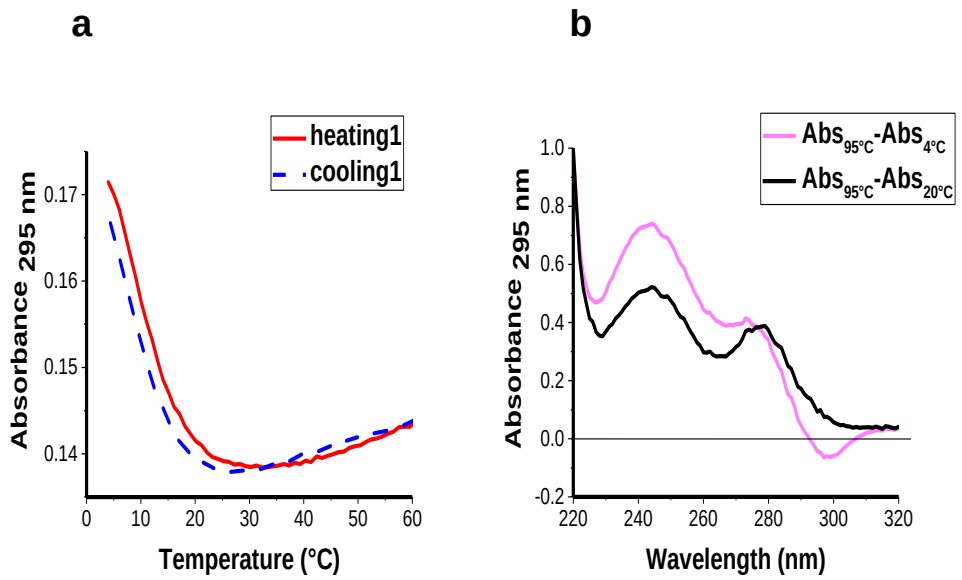


Figure 3.11. UV melting and annealing representative profiles (**a**) and the thermal difference spectra (TDS) (**b**) calculated between 95 and 4°C of 2.5 µM AT5G RNA in KCl (800 mM) and sodium cacodylate buffer (50 mM pH 5.8).

AT5G008230.1	**pH 5.8**	
	Tm /°C	***Ta /°C***
100 mM	16.8 ± 0.6	15.8 ± 1.5
PEG 8,000 (20%)	23.9 ± 5.2	28.2 ± 4.6
800 mM	7.3 ± 1.0	*nd

Table 7. UV-melting and annealing profile of AT5G in sodium cacodylate buffer (50 mM pH 5.8) in KCl (100 mM, 800 mM), and PEG 8,000 (20%)comparing the different T_m , T_a ,average and standard deviation under the same 2.5 µM AT5G RNA conditions.

***nd** no discernable**.**

We concluded from our last results that the PEG was changing the pH of the samples, and the melting temperatures were continue increasing after every

heating cycle. A new control experiment with pH was performed using the heating block to mimic the 3 heating/cooling cycles from the UV machine. Our aim was to observe if pH of the buffers was changed after the samples were exposed to the heating cycles using the heating block. Samples were prepared with the same buffer conditions and PEG 8,000 at 40%. For the samples used for control with 20% and 40% PEG 8,000, pH was measured before and after the PEG was added. We have observed a change on pH for both samples with PEG, and the sample PEG 8,000 (40%) having the highest change of pH or *Δ*pH of 0.26 units. Then, after the first heating cycle, we have found for the No-PEG sample the pH has decreased in -0.05 units, then, after the same sample was expose to the 3 cycles, pH has decreased in 0.03 units. For the sample with PEG 8,000 (20%) the pH has decreased in 0.26 units through the 3 cycles after the PEG was added. We have observed sample with PEG 8,000 (40%) has showed the highest decreased in the pH in first cycle started with pH 5.5 and after cycle 3 the pH was 4.98 (Table 8).

	pH 5.8					
	No PEG	***Δ*pH**	**PEG 20%**	***Δ*pH**	**PEG 40%**	***Δ*pH**
"NO cycle"	5.56		*before/after 5.64/5.95	-0.05	*before/after 5.64/5.9	-0.26
Cycle 1	5.51	-0.05	5.72	-0.23	5.5	-0.4
Cycle 2	5.52	-0.04	5.71	-0.24	5.25	-0.65
Cycle 3	5.53	-0.03	5.69	-0.26	4.98	-0.92

Table 8. Showing the delta of pH (*Δ*pH) for No-PEG, PEG 20% and PEG 40% potassium chloride (**100 mM**) and sodium cacodylate (50 mM pH 5.8) using the heating block.

*before/after the PEG 8,000 (20% or 40%) was added to the control sample.

Turning now to the second set for this part of control experiments which have the same conditions, but with 800 mM KCl. The protocol to follow was the same for the last set of this experiments. The pH changed for both samples with PEG 8,000 at 40% showing the highest difference or *Δ*pH of 0.42 units. Overall, we have measured the *Δ*pH for No-PEG, 20% PEG and 40% PEG samples after they were exposed to the three heating/cooling cycles, and we have observed a decrease of the pH. The sample with 40% PEG showed the higher decreased or *Δ*pH of 0.57 units (Table 9).

	pH 5.8					
	No PEG	***Δ*pH**	**PEG 20%**	***Δ*pH**	**PEG 40%**	***Δ*pH**
"NO cycle"	5.57		*before/after 5.66/5.84	0.18	*before/after 5.84/6.26	0.42
Cycle 1	5.52	-0.05	5.72	-0.12	6.05	-0.21
Cycle 2	5.53	-0.04	5.69	-0.15	5.95	-0.31
Cycle 3	5.51	-0.06	5.63	-0.21	5.69	-0.57

Table 9. Showing the delta of pH (*Δ*pH) for the set of samples labelled No-PEG, PEG 20% and PEG 40% at **800** mM potassium chloride using the heating block. *before/after the PEG 8,000 (20% or 40%) was added to the control sample.

The results for the *Δ*pH from the last results of control experiments with 100 and 800 mM KCl suggested a change in the pH after our samples were exposed to the heating block. Not only that, but oddly we have found that even though by increasing the concentration of the buffer to 50 mM, we still found a considerable change in the pH. Thus, we looked at the literature and found earlier studies

suggested the presence of different structures and changes in the pH in the buffered solutions. Lewis *et. al.* have reported the formation of a no specific PEG/DNA complex with PEG 8,000 at 30% and 40% (65), and knowing this, Reilly and coworkers tested the effect of the PEG in the pH, and their findings suggested the change in activity is related to the percentage of PEG in the solution and not the size of the PEG (15).

A new stock buffer solution was prepared. But, this time a large stock of 50 mL to make it by measuring the mass of all the components: sodium cacodylate, potassium chloride, and the mass of PEG 8,000 was added according to the desire concentration (20% or 40%). RNAase free water was used to dilute all the components, then pH was adjusted by using HCl or NaOH as appropriate. The pH was measured with a probe for all of them before and after every cycle. Our results for the *Δ*pH for the three samples after the three heating cycles using heating block were reduced in comparison with all control experiments made before. (Table10). Having the *Δ*pH for the sample with 40% PEG of 0.15 units, our findings suggested that using this large stock was safer for our experiments since that way the change on the pH was minimized.

	Samples through 3 cycles pH 5.8					
	No PEG	***Δ*pH**	**PEG 20%**	***Δ*pH**	**PEG 40%**	***Δ*pH**
"NO cycle"	5.78		5.86		5.85	
Cycle 1	5.68	-0.10	5.75	-0.11	5.70	-0.15
Cycle 2	5.74	-0.04	5.81	-0.05	5.78	-0.07
Cycle 3	5.75	-0.03	5.74	-0.12	5.70	**-0.15**

Table 10. Showing the delta of pH (*Δ*pH) for the control experiment obtained from the samples No-PEG, PEG 20% and PEG 40% after they were exposed to the heating block mimicking the heating cycles of the UV.

3.7.2 UV melting/annealing profile with UV with 100 mM KCl, PEG 8,000(20%) and PEG 8,000(40%) using large stock buffer solution.

The UV melting profile for the melting and annealing cycles showed two states, unfolded and folded, for the sample with 100 mM KCl and 50 mM buffer without PEG. The average T_m for the heating/melting curves was found to be 10.7°C ± 1.5 and the T_a for the annealing/ cooling curves is 11°C ± 0.04 (Fig.3.12 a), for the thermal difference spectra TDS at 4°C showed i-motif structure with a positive peak at 243 nm and a second peak around 297 nm (Fig.3.12 b).

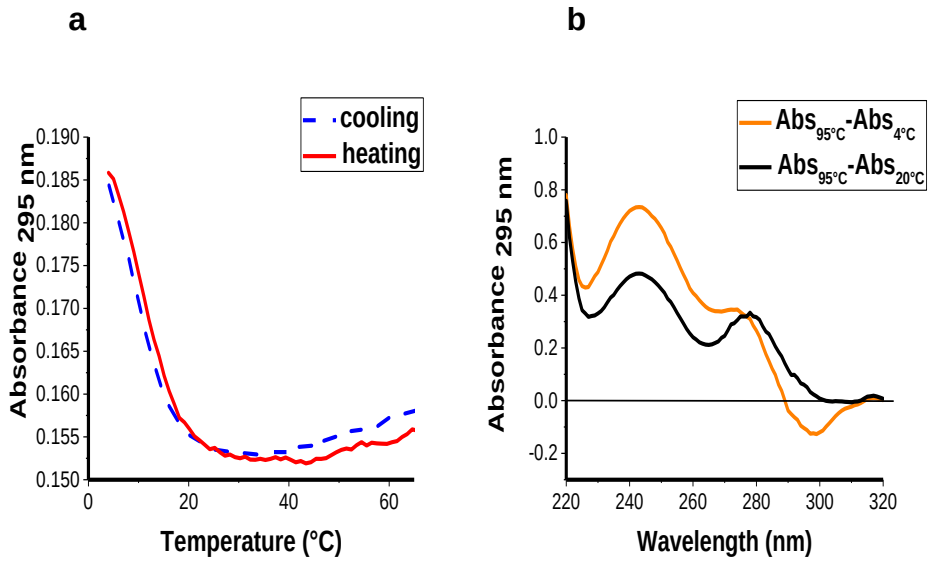


Figure 3.12. UV melting and annealing representative profiles (**a**) and the thermal difference spectra (**b)** calculated between 95 and 4°C of AT5G of 2.5 µM AT5G RNA in KCl (100 mM) and sodium cacodylate buffer (50 mM pH 5.8).

The thermal profile for the melting and annealing curves for the PEG 8,000 (20%) showed an increase of 6°C in comparison with the No-PEG sample showing already some stabilization, and having an average for T_m is 16.8 ± 0.6 and T_a is 14.8°C ± 0.6 (Fig.3.13 a and Table11). The thermal difference spectra at 4°C showed also i-motif structure with a major peak around 243 nm and a second peak around 297 nm (Fig.3.13 b).

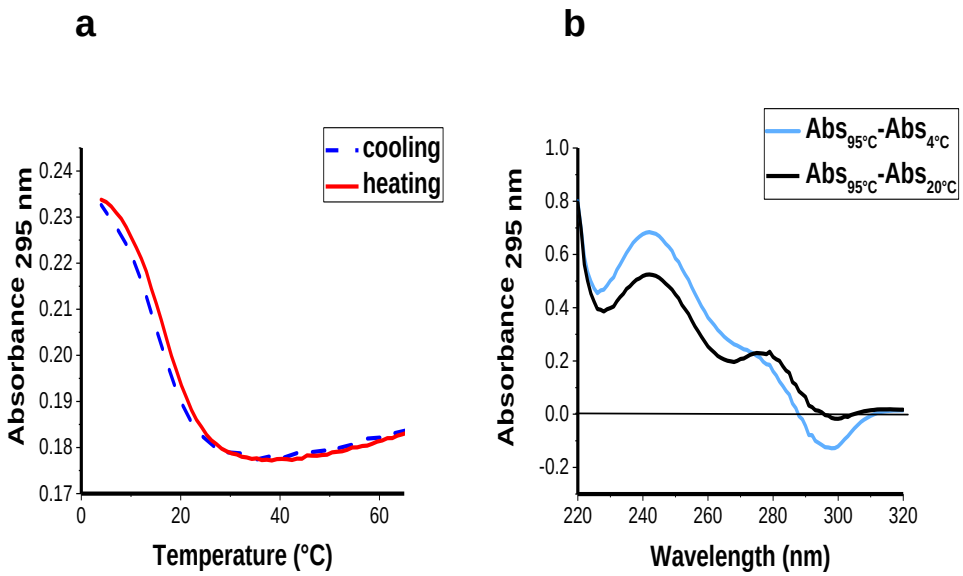


Figure 3.13. UV melting and annealing profiles (**a**) and the thermal difference spectra (**b)** calculated between 95 and 4°C of AT5G of 2.5 µM AT5G RNA in KCl (100 mM), sodium cacodylate buffer (50 mM pH 5.8) and PEG 8,000(20%).

We have observed an increase of the thermal stability as the percentage of the PEG 8,000 increased to 40% **(**Fig.3.14 a) The T_m and T_a both are closer to 20°C, and after established the presence of the secondary structure, the thermal difference spectra was used to indicate the dominant folded species. The TDS curve at 4°C has showed a classic i-motif structure with a major positive peak around 243 nm and the second peak at 296 nm (Fig.3.14 b).

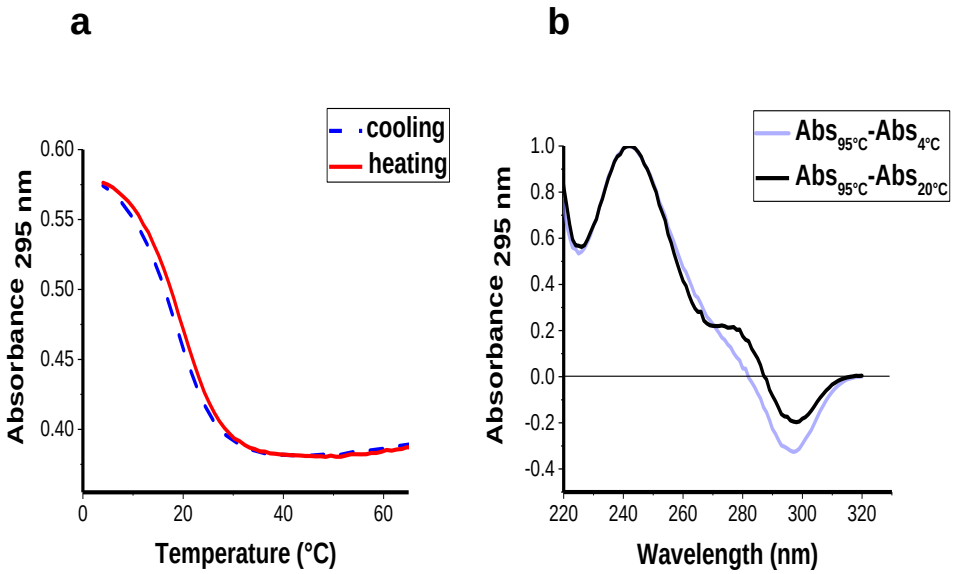


Figure 3.14 UV melting and annealing profiles (**a**) and the thermal difference spectra (**b)** calculated between 95 and 4°C of AT5G of 2.5 µM AT5G RNA in KCl (100 mM), sodium cacodylate buffer (50 mM pH 5.8) and PEG 8,000 (40%).

AT5G008230.1	**pH 5.8**	
	***T_m* /°C**	***T_a* /°C**
No-PEG	10.74 ± 1.5	11.1 ± 0.04
PEG 8,000 (20%)	16.8 1± 0.6	14.8 ± 0.6
PEG 8,000 (40%)	19.17 ± 1.0	18.19 ± 1.0

Table 11. UV-melting and annealing profile of AT5G in KCl (10 mM), sodium cacodylate buffer (50 mM pH 5.8) after control experiments part III comparing the different T_m, T_a ,average and standard deviation under the same RNA conditions for No-PEG, PEG 8,000 (20%) and PEG 8,000 (40%).

3.8 Mimicking crowding conditions and characterized with Circular Dichroism using large stock buffer solution and KCl (100 and 800 mM), PEG 8,000 (20%, 40%).

3.8.1 Introduction

CD was used to further characterize the folded RNA species involved in the UV melting and annealing profile. To mimic the crowding conditions of the cell, we analyzed the sequences in the presence and absence of 20% and 40% PEG 8,000, and two different concentrations of 100 and 800 mM potassium chloride at room temperature.

3.8.2 Results and Discussion

CD signal was recorded for both samples without PEG with 100 and 800 mM KCl showed a positive peak around 280 nm and a negative peak around 240 nm that is within the limits of normal A-form for RNA according to Sugimoto (139). For the samples at 20% PEG, CD signal showed the highest peak at 280 nm suggesting the PEG 8,000 stabilized the oligo already

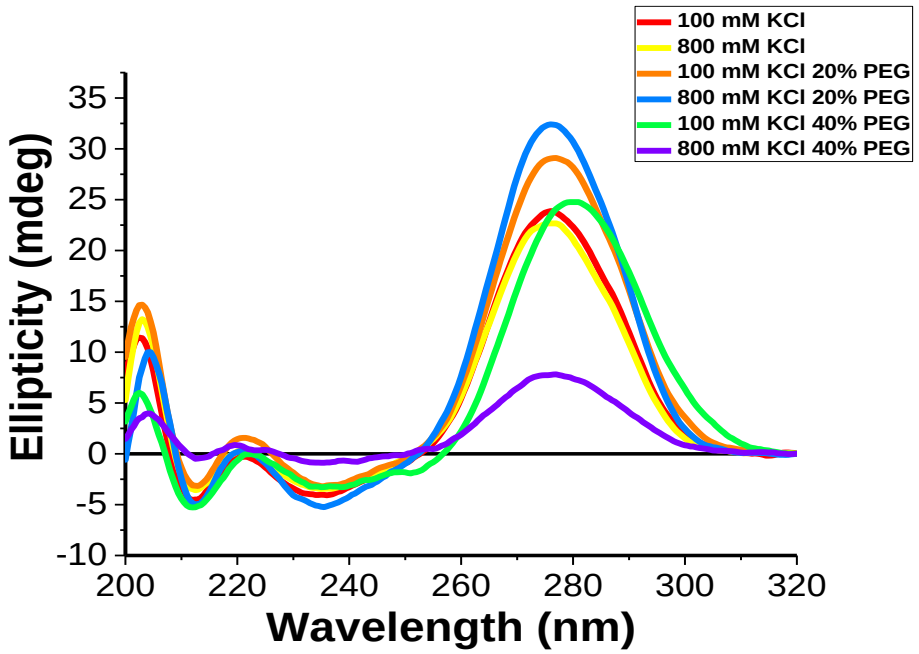


Figure 3.15 Circular Dichroism of 10 µM AT5G in the absence and present PEG 8,000 at 20% and 40%, sodium cacodylate (10 mM pH 5.8.) potassium chloride (100 and 800 mM).

3.8.3 Conclusions

For the samples at 800 mM potassium, while the sample at PEG 8,000 (20%) showed the biggest dominant peak around 280 nm than any of the samples, the sample with PEG 8,000 (40%) showed the shortest peak with a small shift to the lower wavelength. Samples with 100 mM potassium chloride and PEG 8,000 (40%) showed a dominant peak closer to 288 nm and a negative peak around 255 nm confirming the presence of i-motif structure, and that the addition of PEG 8,000 at 40% stabilizes i-motif structure (105) (Fig.3.15).

3.9 Circular Dichroism using PEG 8,000 (40%) at room temperature and 4° C.

3.9.1 Introduction

From our last CD experiments trying to characterize further the folded RNA species, we found the sample with PEG 8,000 (40%) and 800 mM potassium chloride at room temperature showed CD spectrum within the limits of i-motif structure. We were expecting some changes in the CD signal due to high concentration of potassium chloride which probably destabilized i-motif structure. Additionally, the UV-melting profile for the experiments at this concentration made with the same stock of buffer showed an average of melting temperature around 20°C at room temperature. Keeping this in mind, we conducted experiments at 4°C since the melting cycles for our UV experiments started at 4°C.

3.9.2 Results and Discussion

We have found both samples with 800 mM showed a dramatic reduced positive peak compared with the samples at room temperature. For the samples at 4°C, both samples with 800 mM and 100 mM showed CD spectrum within the limits of i-motif with a positive peak at 288 nm and a negative peak at 255 (105).

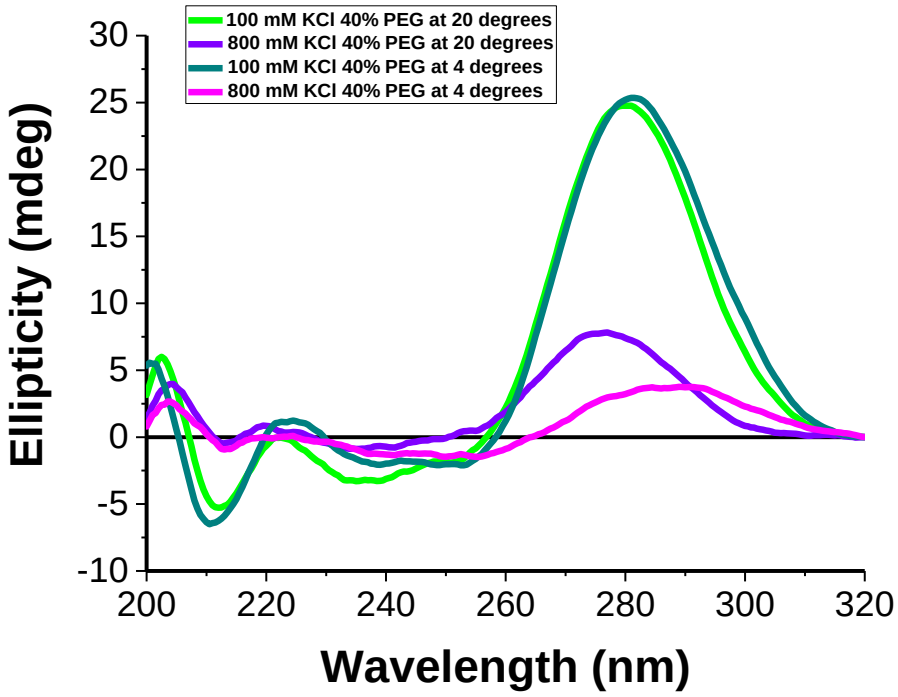


Figure 3.16 Circular Dichroism spectra of 10 µM AT5G with PEG 8,000 (40%) sodium cacodylate (10 mM pH 5.8.) potassium chloride (100 and 800 mM) at 4°C and 20°C.

3.9.3 Conclusions

We were expecting at this temperature more i-motif formation knowing that i-motif structure tends to stabilize at lower temperatures (16). But, we did not observe this, suggesting that the lower temperature has an effect in the solubility of the samples. For the samples at 20°C, while sample with 100 mM showed a CD spectrum within the limits of i-motif, the sample with 800 mM shifted to the lower wavelength with a positive peak around the 280 nm, and CD signal has reduced suggesting less thermal stability (Fig.3.16).

Chapter 4: Conclusions and Future work

4.1 Overall conclusions.

This thesis has described the study of RNA i-motifs in the library of sequences $[C_nU_3]_3U_n$ and a plant sequence AT5G08230.1 RNA identified in *Arabidopsis Thaliana*.

In chapter 2, we have studied and characterized an RNA i-motif version of the sequence $[C_nU_3]_3U_n$ at two different pH: 5.5 where DNA i-motif has been observed stable, and at physiological pH. On one hand, at pH 5.5, while UV and TDS results have found a minimum of two cytosines for tract length to form i-motif structure, CD spectroscopy results performed at room temperature have observed minimum of three cytosines per tract to fold into i-motif structure at 288 nm. From our analysis made with the CD results maxima ellipticity, ellipticity at 200 nm and ellipticity at 288 nm, we have found the pH_T reached a plateau in sequences with 5, 6 or more cytosines suggesting pH stability. Overall, these findings might suggest sequences with five cytosines or more per tract contribute to form stable i-motif at pH 5.5. Comparing the C_nT_3 and C_nU_3 thermal profiles with UV spectroscopy, we found the C_nT_3 have shown higher thermal stability than C_nU_3. The T_m and T_a versus the number of cytosines per tract length was analyzed also, and we have concluded the $\boldsymbol{T_m}$ and $\boldsymbol{T_a}$ increased with the number of cytosines increase. Our findings using TDS for both C_nT_3 and C_nU_3 i-motif was the dominant secondary structure for all the sequences except C_1T_3 and C_1U_3. The level of hysteresis was analyzed and this degree has reached an upper limit in sequences with 6 or more cytosines per tract length. Overall, the RNA sequence C_nU_3 has shown a higher degree of hysteresis suggesting less thermal stability than the DNA sequence C_nT_3. While at pH 7.4 we could not confirm the presence of RNA i-motif structure from any of the cytosine-rich sequences from our library using UV, TDS or CD, the sequence C_nT_3 have found i-motif structure at pH 7.4 without using a molecular crowder. In our case, we might suggest the use of molecular crowders of different sizes and percentages.

In chapter 3, the first example of characterization of an RNA i-motif forming sequence from plants have been described using the sequence AT5G08230.1 RNA. Our analysis of the sequence using CD spectroscopy at different pH from 4 to 8 including 5.8 suggested this sequence was able to fold into i-motif at the most

acidic pH confirming the preference of i-motif for the acidic environment. Samples at pH 4, 4.5 and 5 have shown the highest CD signal intensity at 288 nm where i-motif folds. Samples at pH 4, 4.5 and 5 revealed a positive peak around 288 and a second peak at 255 nm confirm the presence of i-motif structure. For the samples at basic pH, they were unfolded at 288 nm. But, they showed a positive peak around 275 nm and negative peak around 215 nm which would correspond to an unfolded conformation or another structure but not i-motif. We found the CD signal shifted up at 200 nm, and the values for the ellipticity at 200 nm were plotted versus the pH, samples with the pH 6, 6.5 and 7 showed the highest CD signal. CD-titration with KCl was performed looking for any conformational changes of our sequences, we have observed the sample without any KCl started like a random coil with a positive peak around the 278 nm after that they were some small changes but not dramatic in the CD signal. For example, from 278 nm for the sample without any addition of KCl to the last addition showing a positive peak around 275 nm.

The results from UV spectroscopy at pH 5.8 revealed at 100 mM the T_m and T_a closer to 12°C and 800 mM both T_m and T_a have shown closer to 6°C suggesting a lower thermal stability in the presence of high concentration of KCl. On the other hand, when PEG 8,000 at 40% was used, UV spectroscopy showed an increase of the T_m of 19.17 ± 1.0 and T_a of 18.19 ± 1.0 suggesting a higher thermal stability. TDS results revealed a positive peak around the 240 nm and a second peak around 297 nm that correspond to the limits of a classical i-motif structure. CD results confirmed the presence of i-motif structure with a positive peak around the 288nm and a second peak at 255 nm. We might conclude RNA i- motif can be stabilized by molecular crowding using the PEG 8,000 (40%).

After the control experiments, it was found that using the correct kind of buffer was essential otherwise we have observed the pH has changed. For the sample without any PEG the average temperature for T_m 10.74°C and T_a 11.1°C, at 20% PEG 8,000 there was an increase in the melting/annealing profiles T_m was 16.8 °C and T_a 14.8°C suggesting the PEG was stabilizing the structure. And, for the sample at 40% PEG T_m was 19.17°C and T_a 18.9°C showing the higher thermal stability of all the 3 samples. Once again, the presence of a secondary structure

was confirmed with UV spectroscopy further characterization was made with CD spectroscopy to see what was the secondary structure involved. At 4°C, both samples with 40% PEG 8,000 at 100 and 800 mM folded into i-motif structure showing a positive peak at 288 nm and a second peak at 255. At 20°C, the sample with 40% PEG 8,000 at 100 mM and 800 mM showed i-motif structure. Overall, our results demonstrated PEG 8,000 40% had the better effect stabilizing i-motif structure at pH 5.8 in AT5G08230.1 RNA. Moreover, our data indicated that the AT5G08230.1 RNA fold into i-motif at acidic pH 5.8 where plants normally thrive. To our knowledge, this is the first report of RNA i-motif structure in plants a result that cast a new light on investigations on RNA i-motif in plants. This research will have impact on the understanding of i-motifs and nucleic acid secondary structure. This will improve understanding of how plants respond to environmental stress and how to preserve them.

4.2 Future work

For chapter 2, future investigations are necessary to validate the conclusions and findings that were draw from this study of the sequence $[C_nU_3]_3U_n$ including further biophysical characterization of the folded RNA species with PAGE (native polyacrylamide gel electrophoresis) NMR (nuclear magnetic resonance) to develop and confirm with these techniques the i-motif structure. Future research would consider working with different concentrations of the oligo, for example starting from lowest to higher (1.25, 2.5, 5 and 10 mM), different pH 5.5, 6 and 7, and also some molecular crowders like PEG 200, 300, 1,000, 4,000 and 8,000 at different concentrations (20%,30 and 40%). Future research should consider working on our sequence C_5U_3 because our finding confirming sequences with 5 or more cytosines per tract length were more stable. These studies should include to work with this C_5U_3 at pH 6 and 7 with different sizes of polyethylene glycol as a molecular crowder like 200, 300, 1,000, 4,000 and 8,000 at different percentages 20, 40 and 60% including at 4°C where i-motif has showed to be more stable. Our studies could not confirm the RNA i-motif structure at pH 7.4, and future research should include work with a molecular crowders like PEG at different sizes at this specific pH. Future work should consider overcoming the limitations of our research for the sequence $C_{10}U_3$ which was only tested for the CD experiments from pH 4 to 6. And, future work should include to test CD $C_{10}U_3$ using same RNA conditions to start at pH 6.5, and 7, and compare the new results with our data.

For chapter 3, future work for the plant sequence or AT5G would include to continue and further characterized and confirm our findings with other different techniques that we did not use like PAGE to visualize the RNA folded structure, NMR (nuclear magnetic resonance). Our research included testing this plant sequence with different types of environmental stress like acidity stress including pH from 4 to 8, salinity stress for example 100 and 800 mM KCl, but future research should include an intermedium concentration of KCl like 500 or 600 mM to compare, we mimicked drought stress using molecular crowders like PEG 8,000, we should include working on PEG 6,000 because is the polyethylene glycol that have been more widely use in plants. Even though more of this studies in the literature were based on germination test in plants. Future work will develop

a method to stablish a structural characterization of RNA i-motif *in vivo.* Including in vivo assessment of i-motif structural alterations in both mutants and using environmental conditions which induce plant stress response (acidity, salinity and drought). Future work should include to present a new tool for identification of the structure of i-motifs *in vivo* and eventually it can be applied to an organism. Our finding of i-motif was able to fold in a RNA plant sequence is very exciting and future research is needed to confirm this novelty. This is likely to be of interest to those working in the field. Further studies base on these findings will help us to understand plant physiology and eventually to protect crops from drought, salinity, cell damage and stress around the world.

Chapter 5: Materials and Methods

5.1 GENERAL

Chemical reagents were used as they were supplied, and solvents were supplied from Sigma-Aldrich or Fisher.

Oligonucleotides were supplied by Eurogentec (Belgium), synthesized on a 200 nmol scale and purified by HPLC-RP. RNA was dissolved in MilliQ water to give 100 µM. Final concentration was confirmed using by UV spectroscopy using Nanodrop ND-1000 spectrophotometer instrument using the extinction coefficients provided for the sequence by Eurogentec. These stocks were then diluted in the appropriate buffers for further experiments

UV absorption spectroscopy UV melting experiments were performed on a Cary 60 UV-Vis spectrometer (Agilent Technologies) equipped with a TC1 Temperature Controller (Quantum Northwest) and recorded using a low volume masked quartz cuvette (1 cm path length). Oligo was diluted to 2.5 µM in buffer at the desired pH 5.5, 5.8 and 7.4 in samples (200 µl) were transferred to a cuvette, covered with a layer of silicone oil and stoppered to reduce evaporation of the sample.

First Derivative was calculated for every cycle after data was recorded on a Cary 60 UV-Vis spectrometer (Agilent Technologies) equipped with a TC1 Temperature Controller (Quantum Northwest) using a low volume masked quartz cuvette (1 cm path length) by dividing the Δ of absorbance and Δ of the temperature (ΔAbs/ΔT°) and then data obtained was formatted to calculated the top ranked value that will correspond the melting temperature (T_m°) using Microsoft Excel and Origin 2015.

Circular dichroism experiments were recorded on a Jasco J-810 spectropolarimeter under a constant flow of nitrogen using a 1 mm path length quartz cuvette. In a typical experiment, oligonucleotide was diluted in pH 5.5, pH 5.8 or pH 7.4 buffer to a concentration of 10 µM and 200 µL of the RNA solution was scanned at room temperature (20°C) between 200 and 320 nm. Data pitch was set to 0.5 nm and measurements were taken at a scanning speed of 200 nm/min, response time of 1 s, bandwidth of 2 nm and the 100 mdeg sensitivity; each spectrum was the average of three scans. Each curve is the average of 3 scans. Each sample had a scan of the appropriate buffer subtracted from it, the effect of

dilution corrected and was corrected to zero at 320 nm. The data was analysed using Microsoft Excel and Origin 2015.

PDB the protein data base or molecule source archiving three-dimensional structures of proteins, nucleic acids and others biological macromolecules, and others experimental data (140). In the figures in chapter 1, the B-form DNA, the RNA and i-motif structures have each one an specific ID that was obtained from the protein data bank (PDB) available online, and by using UCSF Chimera website the structure was downloaded by fetching with the specific ID, and then adapted and saved as an image.

5.2 FOR CHAPTER 2

UV experiments were carry out in a Cary 60 UV-Vis spectrometer (Agilent Technologies) as it was described in section 5.1. Oligonucleotides from our library of sequences (Table1) were diluted either at pH 5.5 or 7.4 buffer to a concentration of 2.5 µM and 200 µL of sample was scanned. The absorbance of the oligo was measured at 295 nm as the temperature of the sample was held for 10 min at 4°C then heated to 95°C at a rate of 0.5°C per min, held at 95°C before the process was reversed; each melting/annealing process was repeated three times. Data were recorded every 1°C during both melting and annealing and each point was the average of three scans. Data was analyzed using Microsoft Excel and Origin 2015. Melting temperatures (T_m) were determined using the first derivative method. For the thermal difference spectra (TDS) data was calculated by subtracting the spectrum between 220 and 320 nm of the folded structure at 4°C and the unfolded structure at 95°C, then, data was normalized by setting the maximum change in the absorption at +1.

CD experiments were carry out in a Jasco J-810 spectropolarimeter as described in section 5.1. Samples were thermally annealed by heating to 95°C in a heat block and then, allowed to cool down at room temperature overnight. The sequences were annealed at pH 5.5. Oligo was diluted to 10 µM (total volume: 200 µl) in buffer at pH increments of 0.5 pH unit from 4 to 7. The scans were recorded at room temperature (20°C) between 200 and 320 nm. Data pitch was set to 0.5 nm and measurements were taken at a scanning speed of 200 nm/min, response time of 1 s, bandwidth of 2 nm and the 100 mdeg sensitivity; each spectrum was the average of three scans.

Samples containing only buffer were also scanned according to these parameters to allow for blank subtraction. Each sample had a scan of the appropriate buffer subtracted from it, the effect of dilution corrected and was corrected to zero at 320 nm. Transitional pH (**pH_T)** for i-motif was calculated from the inflection point of fitted ellipticity at 200 nm, 288 nm and maxima ellipticity. Final analysis and manipulation of the data was carried out using Excel/Origin 2015.

5.3 FOR CHAPTER 3

CD experiments were carry out in a Jasco J-810 spectropolarimeter as described in section 5.1. Samples were thermally annealed by heating to 95°C in a heat block and then, allowed to cool down at room temperature overnight. Oligo was diluted to 10 µM (total volume: 200 µl) in buffer at pH increments of 0.5 pH unit from 4 to 8 including 5.8 for section 3.2.1, at pH 5.8 with or without PEG 8,000 (40%) for section 3.3.2, and at pH 5.8 at either 100 mM, 800 Mm and PEG 8,000 (20% and 40%) for section 3.6.2. The scans were recorded at room temperature (20°C) between 200 and 320 nm. Data pitch was set to 0.5 nm and measurements were taken at a scanning speed of 200 nm/min, response time of 1 s, bandwidth of 2 nm and the 100 mdeg sensitivity; each spectrum was the average of three scans. Samples containing only buffer were also scanned according to these parameters to allow for blank subtraction. Each sample had a scan of the appropriate buffer subtracted from it, the effect of dilution corrected and was corrected to zero at 320 nm. Transitional pH (**pH_T)** for i-motif was calculated from the inflection point of fitted ellipticity at 200 nm, 275 nm, 288 nm and maxima ellipticity in a sigmoidal curve using Origin.2015. Final analysis and manipulation of the data was carried out using Excel/Origin 2015.

For the CD-**titration** with KCl, the oligo was diluted in buffer pH 5.8 containing 10 mM sodium cacodylate and a stock of 3M potassium chloride was used to add small amounts (1µl) or the equivalent of 17 mM of KCl gradually using a syringe or pipette and the change in the absorbance was monitored. A sample without any addition of KCl was used as a control. After every addition, the CD spectrum was recorded. Transitional pH (**pH_T)** for CD titration with KCl was calculated from the inflection point of fitted ellipticity at 200 nm and maxima ellipticity, and data was analyzed using Microsoft Excel and Origin 2015.

UV experiments were carry out in a Cary 60 UV-Vis spectrometer (Agilent Technologies) as it was described in section 5.1. Oligonucleotide AT5G was diluted at pH 5.8 buffer to a concentration of 2.5 µM either 100 or 800 mM or at 205 and 40% PEG 8,000, and 200 µL of sample was scanned. The absorbance of the oligo was measured at 295 nm as the temperature of the sample was held for 10 min at 4°C then heated to 95°C at a rate of 0.5°C per min, held at 95°C before the process was reversed; each melting/annealing process was repeated three times. Data were recorded every 1°C during both melting and annealing and each point was the average of three scans. Data was analyzed using Microsoft Excel and Origin 2015. Melting temperatures (Tm) were determined using the first derivative method. For the thermal difference spectra (TDS) data was calculated by subtracting the spectrum between 220 and 320 nm of the folded structure at 4°C and the unfolded structure at 95°C, then, data was normalized by setting the maximum change in the absorption at +1.

Control experiments with pH for the first part of the control experiments using the UV spectroscopy machine experiments were carry out Cary 60 UV-Vis spectrometer (Agilent Technologies) as it was described in section 5.1. The pH of the samples was measured using a calibrated probe before and after the PEG 8,000 was added. For the second part of the control experiments using the heating block mimicking the 3 cycles of the UV machine samples were prepared with the following conditions: 50 mM sodium cacodylate, 100 mM KCl, PEG 8,000 at 20% and 40%. Having 3 conditions, 12 samples were prepared including control samples to compare data. The same protocol was taken for the samples at 800 mM KCl. The pH of the sodium cacodylate buffers was measured before using a probe. We also measured the pH for all the samples including control and the 20% and 40% PEG samples before and after adding the cosolutes. For the first cycle, samples were exposed to the heating block to reach 95°C, and held for 10 minutes, then, let them cool down at room temperature or 20°C. After samples were taken out, the pH was measured. Then, samples went to a second heating cycle to reach 95°C, and let them to cool down at room temperature. Then, they were taken out to measure the pH. The same procedure was done for all the samples 3 times, and delta pH was calculated. For the last part of the control experiments the large stock of buffer solution was prepared or 50 mL of No PEG,

20% PEG and 40% PEG, respectively. The pH was measured with a probe for all of the samples before and after every cycle. Then, all the samples were thermally annealed by heating in a heat block at 95°C for 5 minutes and cooled slowly to room temperature overnight. For the first cycle, samples were exposed to the heating block to reach 95°C, and held for 10 minutes, then, let them cool down at room temperature or 20°C, then, after samples were taken out, the pH was measured. Then, samples went to a second heating cycle to reach 95°C, and let them to cool down at room temperature. Then, they were taken out to measure the pH. The same procedure was done for all the samples 3 times, and pH was measured, and delta of pH was calculated.

Chapter 6: References

1. Choi J., Majima T. Conformational changes of non-B DNA. *Chem. Soc. Rev*. **2011**;40(12): 5893-909.
2. Gellert M., Lipsett M. N., Davies D. R. Helix Formation by Guanylic Acid. *Proc. Natl. Acad. Sci. U S A*. **1962**;48(12): 2013-8.
3. Burge S., Parkinson G. N., Hazel P., Todd A. K., Neidle S. Quadruplex DNA: sequence, topology and structure. *Nucleic Acids Res.* **2006**;34(19): 5402-15.
4. Todd A. K., Johnston M., Neidle S. Highly prevalent putative quadruplex sequence motifs in human DNA. *Nucleic Acids Res*. **2005**;33(9): 2901-7.
5. Huppert J. L., Balasubramanian S. Prevalence of quadruplexes in the human genome. *Nucleic Acids Res*. 2005;33(9): 2908-16.
6. Sundquist W. I., Klug A. Telomeric DNA dimerizes by formation of guanine tetrads between hairpin loops. *Nature*. **1989**;342: 825-29.
7. Frank-Kamenetskii M. The turn of the quadruplex? *Nature*. **1989**;342: 737.
8. Huppert J. L., Balasubramanian S. G-quadruplexes in promoters throughout the human genome. *Nucleic Acids Res*. **2007**;35(2): 406-13.
9. Maizels N. Dynamic roles for G4 DNA in the biology of eukaryotic cells. *Nat. Struct. Mol. Biol*. **2006**;13(12): 1055-9.
10. Gehring K., Leroy J. L., Gueron M. A tetrameric DNA structure with protonated cytosine.cytosine base pairs. *Nature*. **1993**;363(6429): 561-5.
11. Mergny J. L., Lacroix, Laurent, Han, Xiaogang, Leroy, J. L. , Helene, Claude. Intramolecular Folding of Pyrimidine Oligodeoxynucleoyides into an i-DNA Motif. *J. Am. Chem. Soc*. **1995**;117(Number 35): 8887-98.
12. Guéron M., Leroy J. L. The i-motif in nucleic acids. *Curr. Opin. Struct. Bio*. **2000**;10(3): 326-31.
13. Wright E. P., Huppert J. L., Waller Z. A. E. Identification of multiple genomic DNA sequences which form i-motif structures at neutral pH. *Nucleic Acids Res.* **2017**;45(6): 2951-9.
14. Gurung S. P., Schwarz C, Hall J. P., Cardin C. J., Brazier J. A. The importance of loop length on the stability of i-motif structures. *Chem. Commun. (Camb)*. **2015**;51(26): 5630-2.
15. Reilly S. M., Morgan R. K., Brooks T. A., Wadkins R.M. Effect of interior loop length on the thermal stability and pK(a) of i-motif DNA. *Biochemistry*. **2015**;54(6): 1364-70.
16. Zhou J, Wei C, Jia G, Wang X, Feng Z, Li C. Formation of i-motif structure at neutral and slightly alkaline pH. *Mol. Biosyst*. **2010**;6(3): 580-6.
17. Leroy J. L., Gueron M, Mergny J-L., Helene C. Intramolecular Folding of a Fragment of the Cytosine-Rich Strand of Telomeric DNA into an I-Motif. *Nucleic Acids Res*. **1994**;22(9): 1600-6.
18. Brooks T.A., Kendrick S, Hurley L. Making sense of G-quadruplex and i-motif functions in oncogene promoters. *FEBS J*. **2010**;277(17): 3459-69.
19. Kendrick S, Akiyama Y, Hecht S. M., Hurley L. H. The i-Motif in the bcl-2 P1 Promoter Forms an Unexpectedly Stable Structure with a Unique 8:5:7 Loop Folding Pattern. *J. Am. Chem. Soc*. **2009**;131(48): 17667-76.
20. Kendrick S, Kang H-J, Alam M.P., Madathil M.M., Agrawal P., Gokhale V., *et al*. The Dynamic Character of the BCL2 Promoter i-Motif Provides a Mechanism for Modulation of Gene Expression by Compounds That Bind Selectively to the Alternative DNA Hairpin Structure. *J. Am. Chem. Soc*. **2014**;136(11): 4161-71.
21. Takahashi S, Brazier J. A., Sugimoto N. Topological impact of noncanonical DNA structures on Klenow fragment of DNA polymerase. *Proc. Nat. Acad. Sci. U. S. A*. **2017**;114(36): 9605-10.

22. Zeraati M., Langley D. B., Schofield P., Moye A. L., Rouet R., Hughes W.E., *et al*. I-motif DNA structures are formed in the nuclei of human cells. *Nature. Chem*. **2018**;10(6): 631-7.
23. Lacroix L., Mergny J. L., Leroy J. L., Helene C. Inability of RNA to form the i-motif: implications for triplex formation. *Biochemistry*. **1996**;35(26): 8715-22.
24. Snoussi K., Nonin-Lecomte S., Leroy J. L. The RNA i-motif. *J. Mol. Biol*. **2001**;309(1): 139-53.
25. Raines C. Accessible research in JXB. *J. Exp. Bot*. **2016**;67(3): 553.
26. Bradshaw, A. D. Some of the evolutionary consequences of being a plant. *Evol. Biol.* **1972**; 525-47.
27. Feder M.E. Plant and Animal Physiological Ecology, Comparative Physiology/Biochemistry, and Evolutionary Physiology: Opportunities for Synergy: An Introduction to the Symposium1. *Integr. Comp. Biol*. **2002**;42(3): 409-14.
28. Day H.A., Pavlou P., Waller Z.A. i-Motif DNA: structure, stability and targeting with ligands. *Bioorg. Med. Chem*. **2014**;22(16): 4407-18.
29. Martinière A., Bassil E., Jublanc E., Alcon C., Reguera M., Sentenac H., *et al*. In Vivo Intracellular pH Measurements in Tobacco and Arabidopsis Reveal an Unexpected pH Gradient in the Endomembrane System. *Plant. Cell*. **2013**;25(10): 4028-43.
30. Ali S., Rizwan M., Qayyum M. F., Ok Y. S., Ibrahim M., Riaz M., *et al*. Biochar soil amendment on alleviation of drought and salt stress in plants: a critical review. *Env. Sci. Pollut. Res*. **2017**;24(14): 12700-12.
31. Wang M., Zheng Q., Shen Q., Guo S. The critical role of potassium in plant stress response. *Int. J. Mol. Sci*. **2013**;14(4): 7370-90.
32. Lager I., Andreasson O., Dunbar. T. L., Andreasson E., Escobar M.A., Rasmusson A.G. Changes in external pH rapidly alter plant gene expression and modulate auxin and elicitor responses. *Plant Cell Env*. **2010**;33(9): 1513-28.
33. Watson J. D., Crick F. H. C. Molecular Structure of Nucleic Acids: A Structure for Deoxyribose Nucleic Acid. *Nature*. **1953**; 171: 737-8.
34. Bowater, R. P. and Waller, Z. A. *DNA Structure*. In: eLS. John Wiley & Sons, Ltd (Ed.). 2014. Available from: doi:10.1002/9780470015902.a0006002.pub2.
35. Mahler H.R., Fraser D. Structural Mutations. *Nature*. **1961**;189: 948-48.
36. Bohdan S., Stephen N., M. B.H. Conformations of the sugar-phosphate backbone in helical DNA crystal structures. *Biopolymers*. **1997**;42(1): 113-24.
37. Sweeney B. P. Watson and Crick 50 years on. From double helix to pharmacogenomics. *Anaesthesia*. **2004**;59(2): 150-65.
38. Watson J. D., Crick F. H. C. Genetical Implications of the Structure of Deoxyribonucleic Acid. *Nature*. **1953**;171: 964-69.
39. Brahms J., Mommaerts W. F. H. M. A study of conformation of nucleic acids in solution by means of circular dichroism. *J. Mol. Bio*. **1964**;10(1): 73-88.
40. Werner G. Extrachromosomal DNA in Bacteria. *Angew. Chem. Int. Ed. Engl*. **1973**;12(7): 517-28.
41. Clark J. *DNA Structure* 2007. Available from: http://www.chemguide.co.uk/organicprops/aminoacids/dna1.html. [Acc.13th Sept 2018].
42. Drew H.R., Wing R.M., Takano T., Broka C., Tanaka S., Itakura K., *et al*. Structure of a B-DNA dodecamer: conformation and dynamics. *Proc. Natl. Acad. Sci. U. S. A*. **1981**;78(4): 2179-83.

43. Jones K.W., Truman D. E. S. A Hypothesis for Deoxyribonucleic Acid Transcription and Messenger Ribonucleic Acid Synthesis in vivo. *Nature*. **1964**;202: 1264-67.
44. Leslie I. A. N. Biochemistry of Heredity: A General Hypothesis. *Nature*. **1961**;189: 260-268.
45. Stanley J. Wendell. M., Bock R. M. Mechanisms of Expression of Genetic Information. *Nature*. **1961**;190: 299-300.
46. Zubay G. A Possible Mechanism for the Initial Transfer of the Genetic Code from Deoxyribonucleic Acid to Ribonucleic Acid. *Nature*. **1958**;182: 112-3.
47. Peselis A., Serganov A. Structural insights into ligand binding and gene expression control by an adenosylcobalamin riboswitch. *Nat. Struct. Mol. Biol*. **2012**;19(11): 1182-4.
48. Crick F. Central Dogma of Molecular Biology. *Nature*. **1970**;227: 561-3.
49. Li G-W., Xie X.S. Central dogma at the single-molecule level in living cells. *Nature*. **2011**;475: 308-15.
50. Commoner B. Deoxyribonucleic Acid and the Molecular Basis of Self-Duplication. *Nature*. **1964**;203: 486-91.
51. Alam M.P. DNA polymerases: structural diversity and common mechanisms. *J. Biol. Chem*. **1999**;274(25): 17395-8.
52. Peterson I. First peek at DNA transcription. *Science News*. **1997**;151(13): 181-8.
53. Pickrell J. K., Marioni J. C., Pai A. A., Degner J. F., Engelhardt B. E., Nkadori E., *et al*. Understanding mechanisms underlying human gene expression variation with RNA sequencing. *Nature*. **2010**;464: 768-72.
54. Lafontaine, D. L. and Tollervey, D. *Ribosomal RNA*. In eLS, John Wiley & Sons, Ltd (Ed.). 2006. Available from: doi:10.1038/npg.els.0003832.
55. Rajat B., Shawn C., Kiley D., Marla G., Mette P-I., Medha R., *et al*. tRNAs: Cellular barcodes for amino acids. *FEBS Let*. **2010**;584(2): 387-95.
56. Klug A. Rosalind Franklin and the Discovery of the Structure of DNA. *Nature*. **1968**;219: 808-44.
57. Arnott S., Chandrasekaran R., Birdsall D. L., Leslie A. G. W., Ratliff R. L. Left-handed DNA helices. *Nature*. **1980**;283: 743-5.
58. Gessner R.V., Frederick C.A., Quigley G.J., Rich A., Wang A.H.J. The Molecular-Structure of the Left-Handed Z-DNA Double Helix at 1.0-a Atomic Resolution - Geometry, Conformation, and Ionic Interactions of D(Cgcgcg). *J. Biol. Chem*. **1989**;264(14): 7921-35.
59. Kang C.H., Berger I., Lockshin C., Ratliff R., Moyzis R., Rich A. Crystal structure of intercalated four-stranded d(C3T) at 1.4 A resolution. *Proc. Natl. Acad. Sci. U. S. A*. **1994**;91(24): 11636-40.
60. Fleming A. M., Ding Y., Rogers R. A., Zhu J., Burton A. D., *et al*. 4n–1 Is a “Sweet Spot” in DNA i-Motif Folding of 2′-Deoxycytidine Homopolymers. *J. Am. Chem. Soc*. **2017**;139(13): 4682-9.
61. Mir B., Serrano I., Buitrago D., Orozco M., Escaja N., González C. Prevalent Sequences in the Human Genome Can Form Mini i-Motif Structures at Physiological pH. *J. Am. Chem. Soc*. **2017**;139(40): 13985-8.
62. Brazier J. A., Shah A., Brown G.D. I-motif formation in gene promoters: unusually stable formation in sequences complementary to known G-quadruplexes. *Chem. Commun. (Camb)*. **2012**;48(87): 10739-41.
63. Sun D., Hurley L. H. The importance of negative superhelicity in inducing the formation of G-quadruplex and i-motif structures in the c-Myc promoter:

implications for drug targeting and control of gene expression. *J. Med Chem.* **2009**;52(9): 2863-74.
64. Rajendran A., Nakano S., Sugimoto N. Molecular crowding of the cosolutes induces an intramolecular i-motif structure of triplet repeat DNA oligomers at neutral pH. *Chem. Commun. (Camb).* **2010**;46(8): 1299-301.
65. Cui J., Waltman P., Le V. H., Lewis E. A. The effect of molecular crowding on the stability of human c-MYC promoter sequence I-motif at neutral pH. *Molecules.* **2013**;18(10): 12751-67.
66. Day H. A., Huguin C., Waller Z. A. E. Silver cations fold i-motif at neutral pH. *Chem. Comm*. **2013**;49(70): 7696-8.
67. Abdelhamid M. A. S., Fábián L., MacDonald C. J., Cheesman M. R., Gates A. J., Waller Z. A. E. Redox-dependent control of i-Motif DNA structure using copper cations. *Nucleic Acids Res*. **2018**;46(12) 5886-93 .
68. Fujii T., Sugimoto N. Loop nucleotides impact the stability of intrastrand i-motif structures at neutral pH. *Phys. Chem. Chemic. Phy*. **2015**;17(26): 16719-22.
69. Liu D., Balasubramanian S. A proton-fuelled DNA nanomachine. *Angew. Chem. Int. Ed. Engl*. **2003**;42(46): 5734-6.
70. Surana S., Bhat J. M., Koushika S.P., Krishnan Y. An autonomous DNA nanomachine maps spatiotemporal pH changes in a multicellular living organism. *Nat. Commun*. **2011**;2: 340-340.
71. Modi S., M. G. S., Goswami D., Gupta G. D., Mayor S., Krishnan Y. A DNA nanomachine that maps spatial and temporal pH changes inside living cells. *Nat. Nanotechnol*. **2009**;4: 325-30.
72. Nesterova I. V., Nesterov E. E. Rational Design of Highly Responsive pH Sensors Based on DNA i-Motif. *J. Am. Chem. Soc*. **2014**;136(25): 8843-6.
73. Peng Y., Wang X., Xiao Y., Feng L., Zhao C., Ren J., *et al*. i-Motif Quadruplex DNA-Based Biosensor for Distinguishing Single- and Multiwalled Carbon Nanotubes. *J. Am. Chem. Soc*. **2009**;131(38): 13813-8.
74. Yang Y., Liu G., Liu H., Li D., Fan C., Liu D.. An Electrochemically Actuated Reversible DNA Switch. *Nano Letters*. **2010**;10(4): 1393-7.
75. Li T., Ackermann D., Hall A. M., Famulok M. Input-dependent induction of oligonucleotide structural motifs for performing molecular logic. *J. Am. Chem. Soc*. **2012**;134(7): 3508-16.
76. Nonin-Lecomte S., Leroy J. L. Structure of a C-rich strand fragment of the human centromeric satellite III: A pH-dependent intercalation topology. *J. Mol. Biol.* **2001**;309: 491-506.
77. Jia-Huai W. Major groove or minor groove? *Nature*. **1986**;319: 183-4.
78. Leroy, J. L., Karim S., Maurice G. Investigation of the energetics of C-H···O hydrogen bonds in the DNA i-motif via the equilibrium between alternative intercalation topologies. *Magn. Reson. Chem*. **2001**;39(S1): 171-6.
79. Zhou T., Chen P., Niu L., Jin J., Liang D., Li Z., *et al*. pH-responsive size-tunable self-assembled DNA dendrimers. *Angew. Chem. Int. Ed. Engl*. **2012**;51(45): 11271-4.
80. Dong Y., Yang Z., Liu D.. DNA nanotechnology based on i-motif structures. *Acc. Chem. Res*. **2014**;47(6): 1853-60.
81. Chen Y., Qu K., Zhao C., Wu L., Ren J., Wang J, *et al*. Insights into the biomedical effects of carboxylated single-wall carbon nanotubes on telomerase and telomeres. *Nat. Comm*. **2012**;3: 1074-1074.

82. Dzatko S., Krafcikova M., Hansel-Hertsch R., Fessl T., Fiala R., Loja T., *et al*. Evaluation of the Stability of DNA i-Motifs in the Nuclei of Living Mammalian Cells. *Angew. Chem. Int. Ed. Engl*. **2018**;57(8): 2165-9.
83. Biffi G., Tannahill D., McCafferty J., Balasubramanian S. Quantitative visualization of DNA G-quadruplex structures in human cells. *Nat. Chemi*. **2013**;5(3): 182-6.
84. Parkinson G. N., Lee M. P. H., Neidle S. Crystal structure of parallel quadruplexes from human telomeric DNA. *Nature*. **2002**;417: 876-80.
85. Balasubramanian S., Hurley L. H., Neidle S. Targeting G-quadruplexes in gene promoters: a novel anticancer strategy? *Nat. Rev. Drug Discov*. **2011**;10(4): 261-75.
86. Sun H., Xiang J., Shi Y., Yang Q., Guan A., Li Q., *et al*. A newly identified G-quadruplex as a potential target regulating Bcl-2 expression. *Biochim. Biophys. Acta*. **2014**;1840(10): 3052-7.
87. Simonsson T., Pribylova M., Vorlickova M. A nuclease hypersensitive element in the human c-myc promoter adopts several distinct i-tetraplex structures. *Biochem. Bioph. Res. Co.* **2000**;278(1): 158-66
88. Fernando H., Reszka A. P., Huppert J., Ladame S., Rankin S., Venkitaraman A. R., *et al*. A conserved quadruplex motif located in a transcription activation site of the human c-kit oncogene. *Biochemistry*. **2006**;45(25): 7854-60.
89. Collin D., Gehring K. Stability of chimeric DNA/RNA cytosine tetrads: Implications for i-motif formation by RNA. *J. Am. Chem. Soc*. **1998**;120(17): 4069-72.
90. Robidoux S., Damha M. J. D-2-Deoxyribose and D-Arabinose, but not D-Ribose, Stabilize the Cytosine Tetrad (i-DNA) Structure. *J. Biom. Struct. Dynam*. **1997**;15(3): 529-35.
91. Bradbury E. M., Crane-Robinson C. High Resolution Nuclear Magnetic Resonance Studies of Biopolymers. *Nature*. **1968**;220: 1079-84.
92. Han X., Leroy J. L., Gueron M. An intramolecular i-motif: the solution structure and base-pair opening kinetics of d(5mCCT3CCT3ACCT3CC). *J. Mol Biol*. **1998**;278(5): 949-65.
93. Giessen B. C., Gordon G. E. X-ray Diffraction: New High-Speed Technique Based on X-ray Spectrography. *Science*. **1968**;159(3818): 973-5.
94. Marsh R. E., Bierstedt R., Eichhorn E. L. The crystal structure of cytosine-5-acetic acid. *Acta Crystallographica*. **1962**;15(4): 310-6.
95. Langridge R., Rich A. Molecular Structure of Helical Polycytidylic Acid. *Nature*. **1963**;198: 725-8.
96. Chirio-Lebrun M-C., Prats M. Fluorescence resonance energy transfer (FRET): Theory and experiments. *Biochem. Educ*. **1998**;26(4): 320-3.
97. Mergny J-L. Fluorescence Energy Transfer as a Probe for Tetraplex Formation: The i-Motif. *Biochemistry*. **1999**;38(5): 1573-81.
98. Mergny J-L., Lacroix L. Analysis of thermal melting curves. *Oligonucleotides*. **2003**;13(6): 515-37.
99. Smyth M. S., Martin J. H. J. x Ray crystallography. *Mol. Pathol*. **2000**;53(1): 8-14.
100. Chen L., Cai L., Zhang X., Rich A. Crystal structure of a four-stranded intercalated DNA: d(C4). *Biochemistry*. **1994**;33(46): 13540-6.
101. Berger I., Cai L., Chen L., Rich A. Energetics of the lattice: Packing elements in crystals of four-stranded intercalated cytosine-rich DNA molecules. *Biopolymers.* **1997**;44(3): 257-67.

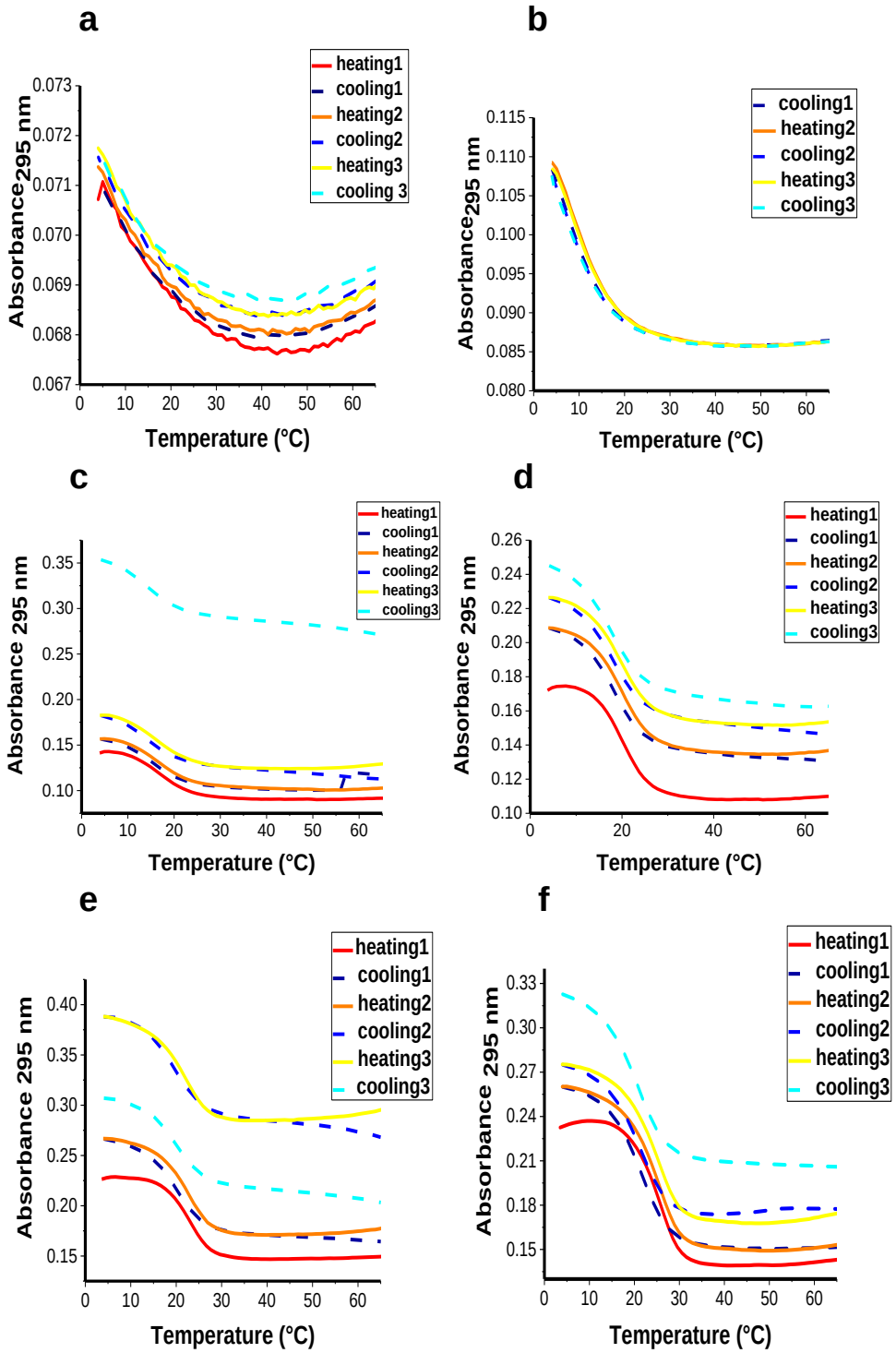
102. Murat P., Singh Y., Defrancq E. Methods for investigating G-quadruplex DNA/ligand interactions. *Chem. Soc. Rev*. **2011**;40(11): 5293-307.
103. Hellman L. M., Fried M. G. Electrophoretic mobility shift assay (EMSA) for detecting protein–nucleic acid interactions. *Nat. Protoc***. 2007**;2: 1849-61.
104. Mergny JL, Li J, Lacroix L, Amrane S, Chaires JB. Thermal difference spectra: a specific signature for nucleic acid structures. *Nucleic Acids Res*. **2005**;33(16): e138.
105. Kypr J., Kejnovska I., Renciuk D., Vorlickova M. Circular dichroism and conformational polymorphism of DNA. *Nucleic Acids Res*. **2009**;37(6): 1713-25.
106. Huey R.B., Carlson M., Crozier L., Frazier M., Hamilton H., Harley C., *et al*. Plants versus animals: do they deal with stress in different ways? *Integr. Comp. Biol*. **2002**;42(3): 415-23.
107. Bose J., Babourina O., Shabala S., Rengel Z. Aluminum-dependent dynamics of ion transport in Arabidopsis: specificity of low pH and aluminum responses. *Physiol Plant*. **2010**;139(4): 401-12.
108. Bodner G., Nakhforoosh A., Kaul H-P. Management of crop water under drought: a review. *Agr. Sust. Devel*. **2015**;35(2): 401-42.
109. Larcher W., Heber U., Santarius K. A. *Limiting Temperatures for Life Functions*. Berlin, Heidelberg: Springer; 1973. p. 195-263.
110. Gupta B., Huang B. Mechanism of Salinity Tolerance in Plants: Physiological, Biochemical, and Molecular Characterization. *Int. J. Genom*. **2014**;2014: 18 pages.
111. Youssef R. A., Chino M. Root-induced changes in the rhizosphere of plants. I. pH changes in relation to the bulk soil. *Soil Sci. Plant. Nutr*. **1989**;35(3): 461-8.
112. Hinsinger P., Plassard C., Tang C., Jaillard B. Origins of root-mediated pH changes in the rhizosphere and their responses to environmental constraints: A review. *Plant Soil*. **2003**;248(1): 43-59.
113. Ecker J. R., Davis R. W. Inhibition of gene expression in plant cells by expression of antisense RNA. *Proc. Natl. Acad. Sci. U. S. A*. **1986**;83(15): 5372-6.
114. Hammond-Kosack K. E, Jones J. D. Resistance gene-dependent plant defense responses. *Plant Cell*. **1996**;8(10): 1773-91.
115. Gao D., Knight M.R., Trewavas A.J., Sattelmacher B., Plieth C. Self-reporting Arabidopsis expressing pH and [$Ca2^{+}$] indicators unveil ion dynamics in the cytoplasm and in the apoplast under abiotic stress. *Plant Physiol*. **2004**;134(3): 898-908.
116. Ellis R.J. Macromolecular crowding: an important but neglected aspect of the intracellular environment. *Curr. Opin. Struct. Biol.* **2001**;11(1): 114-9.
117. Zimmerman S.B., Trach S.O. Estimation of macromolecule concentrations and excluded volume effects for the cytoplasm of Escherichia coli. *J. Mol. Biol*. **1991**;222(3): 599-620.
118. Ellis R. J. Macromolecular crowding: obvious but underappreciated. *Trends Biochem. Sci*. **2001**;26(10): 597-604.
119. Mouillon J. M., Eriksson S. K., Harryson P. Mimicking the plant cell interior under water stress by macromolecular crowding: disordered dehydrin proteins are highly resistant to structural collapse. *Plant Physiol*. **2008**;148(4): 1925-37.
120. Miyoshi D., Sugimoto N. Molecular crowding effects on structure and stability of DNA. *Biochimie*. **2008**;90(7): 1040-51.

121. Minton A.P. The influence of macromolecular crowding and macromolecular confinement on biochemical reactions in physiological media. *J. Biol. Chem*. **2001**;276(14): 10577-80.
122. Kilburn D., Roh J. H., Guo L., Briber R. M, Woodson S. A. Molecular crowding stabilizes folded RNA structure by the excluded volume effect. *J. Am. Chem. Soc.* **2010**;132(25): 8690-6.
123. Strulson C. A. , Boyer J. A. , Whitman E. E., Bevilacqua P. C. Molecular crowders and cosolutes promote folding cooperativity of RNA under physiological ionic conditions. *RNA*. **2014**;20(3): 331-47.
124. Zhao C., Ren J., Qu X. Single-walled carbon nanotubes binding to human telomeric i-motif DNA under molecular-crowding conditions: more water molecules released. *Chemistry*. **2008**;14(18): 5435-9.
125. Banks D. S., Fradin C. Anomalous diffusion of proteins due to molecular crowding. *Biophys. J*. **2005**;89(5): 2960-71.
126. Minton A. P. Macromolecular crowding and molecular recognition. *J. Mol. Recognit*. **1993**;6(4): 211-4.
127. Bhavsar-Jog Y. P., Van Dornshuld E., Brooks T. A., Tschumper G. S., Wadkins R. M. Epigenetic modification, dehydration, and molecular crowding effects on the thermodynamics of i-motif structure formation from C-rich DNA. *Biochemistry*. **2014**;53(10): 1586-94.
128. Tyrrell J., Weeks K. M., Pielak G. J. Challenge of mimicking the influences of the cellular environment on RNA structure by PEG-induced macromolecular crowding. *Biochemistry*. **2015**;54(42): 6447-53.
129. Allen F. S., Gray D. M., Roberts G. P., Tinoco I., Jr. The ultraviolet circular dichroism of some natural DNAs and an analysis of the spectra for sequence information. *Biopolymers*. **1972**;11(4): 853-79.
130. Wells B. D., Yang J. T. A computer probe of the circular dichroic bands of nucleic acids in the ultraviolet region. I. Transfer ribonucleic acid. *Biochemistry*. **1974**;13(7): 1311-6.
131. El Rabey H. A., Al-Malki A. L., Abulnaja K. O. Proteome Analysis of Date Palm (Phoenix dactylifera L.) under Severe Drought and Salt Stress. *Int. J. Genomics*. **2016**; 8 pages.
132. Ding Y., Tang Y., Kwok C. K., Zhang Y., Bevilacqua P. C, Assmann S. M. In vivo genome-wide profiling of RNA secondary structure reveals novel regulatory features. *Nature*. **2013**;505: 696-700.
133. Jali S. S., Rosloski S. M., Janakirama P., Steffen J. G., Zhurov V., Berleth T., *et al*. A plant-specific HUA2-LIKE (HULK) gene family in *Arabidopsis thaliana* is essential for development. *Plant J*. **2014**;80(2): 242-54.
134. Rodríguez-Cazorla E., Ripoll J. J., Andújar A., Bailey L. J., Martínez-Laborda A., Yanofsky M.F., *et al*. K-homology Nuclear Ribonucleoproteins Regulate Floral Organ Identity and Determinacy in Arabidopsis. *PLOS Gen*. **2015**;11(2): e1004983.
135. Shabala S, Pottosin I. *Potassium and Potassium-Permeable Channels in Plant Salt Tolerance*. 2010. [87-110]. Available from: doi: 10.1007/978-3-642-10494-7_5
136. Fabian C., Reimann C., Fabian K., Birke M., Baritz R., Haslinger E. GEMAS: Spatial distribution of the pH of European agricultural and grazing land soil. *App. Geochem*. **2014**;48: 207-16.

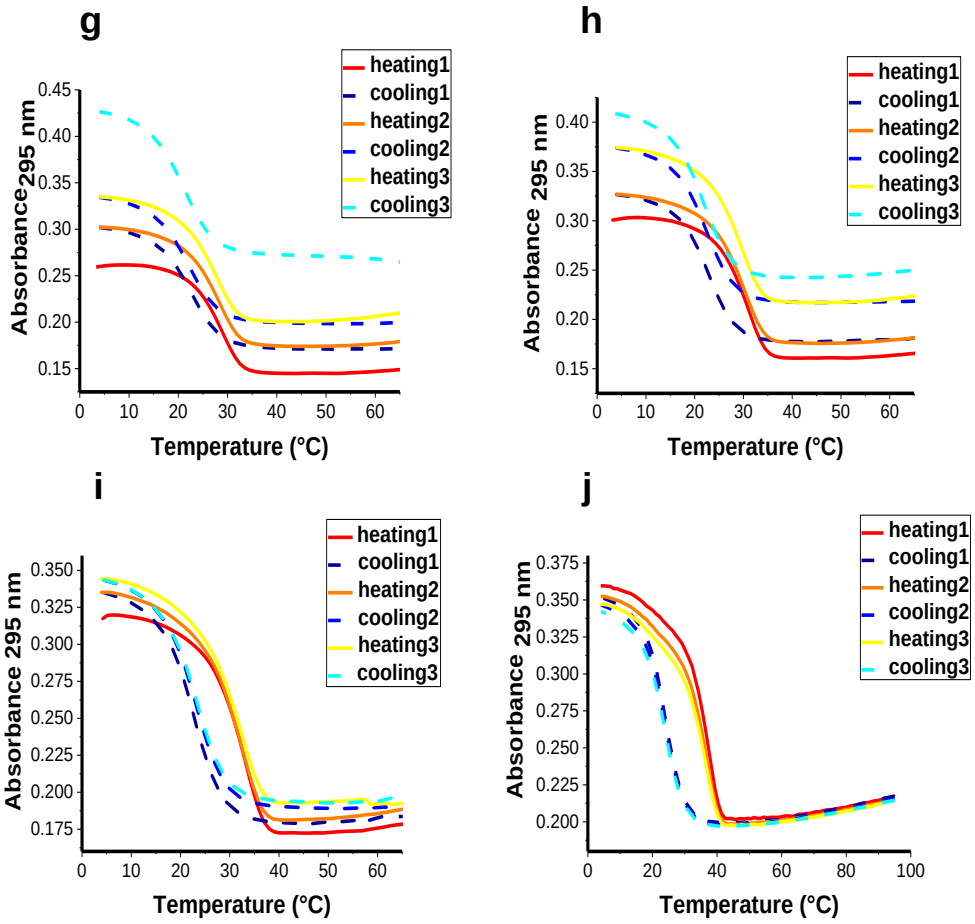
137. Arnon D. I., Johnson C. M. Influence of Hydrogen Ion Concentration on the Growth of Higher Plants under Controlled Conditions. *Plant Physiol*. **1942**;17(4): 525-39.
138. Arakawa T., Timasheff S. N. Mechanism of poly(ethylene glycol) interaction with proteins. *Biochemistry*. **1985**;24(24): 6756-62.
139. Sugimoto N., Nakano S., Katoh M., Matsumura A., Nakamuta H., Ohmichi T., *et al*. Thermodynamic parameters to predict stability of RNA/DNA hybrid duplexes. *Biochemistry*. **1995**;34(35): 11211-6.
140. Burley S. K., Berman H. M., Kleywegt G. J., Markley J. L., Nakamura H., Velankar S. *Protein Data Bank (PDB): The Single Global Macromolecular Structure Archive*. In: Protein Crystallography 2017 (pp 627-41). Humana Press, New York, NY.

Appendix

A1. UV melting and annealing profiles and the thermal difference spectra for $C_n U_3$ at pH 5.5

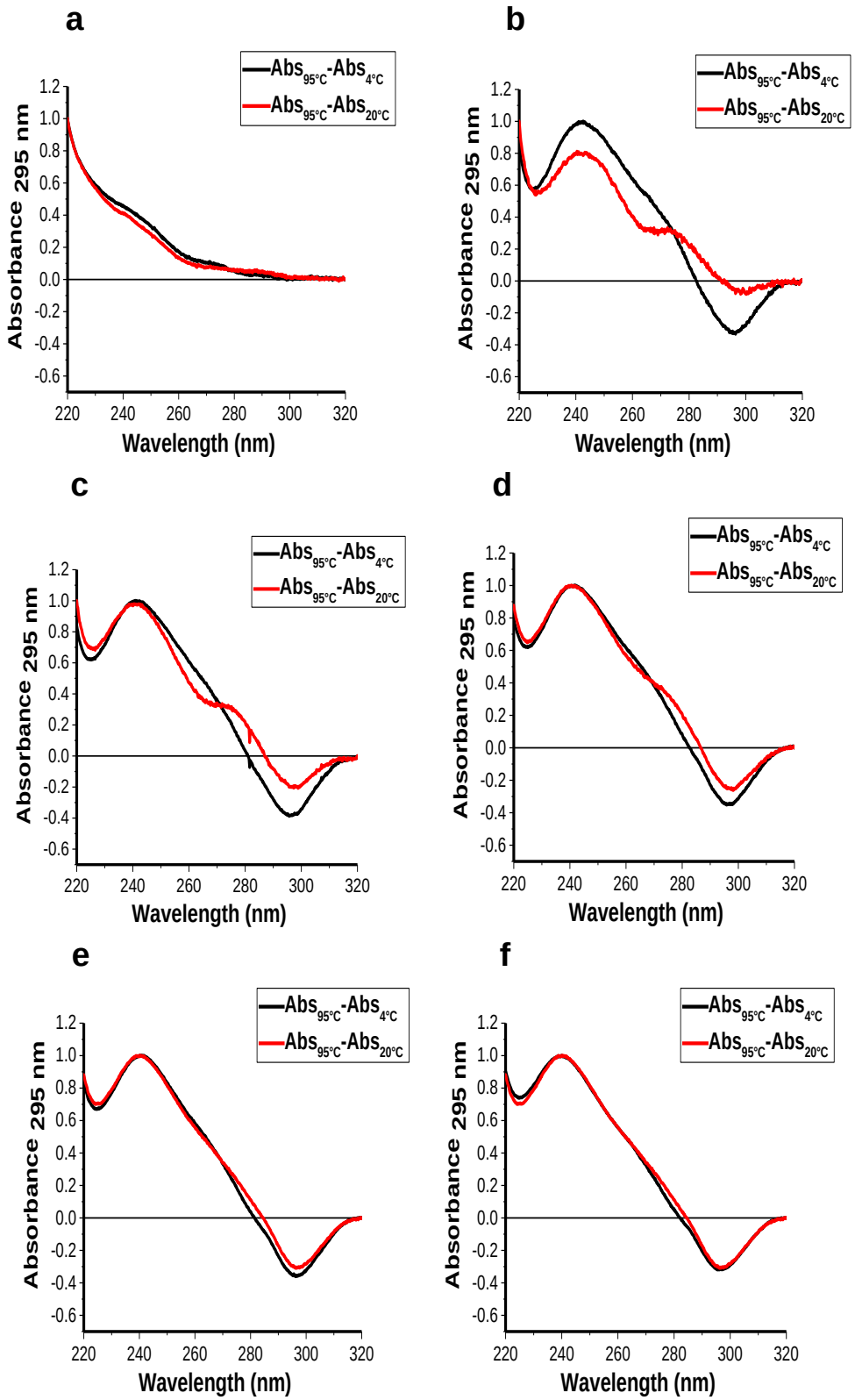


UV melting/annealing profiles of 2.5µM RNA in sodium cacodylate buffer (10 mM pH 5.5) and NaCl (100 mM) (**a**)C_1U_3, (**b**)C_2U_3, (**c**)C_3U_3, (**d**)C_4U_3, (**e**)C_5U_3, (**f**)C_6U_3.

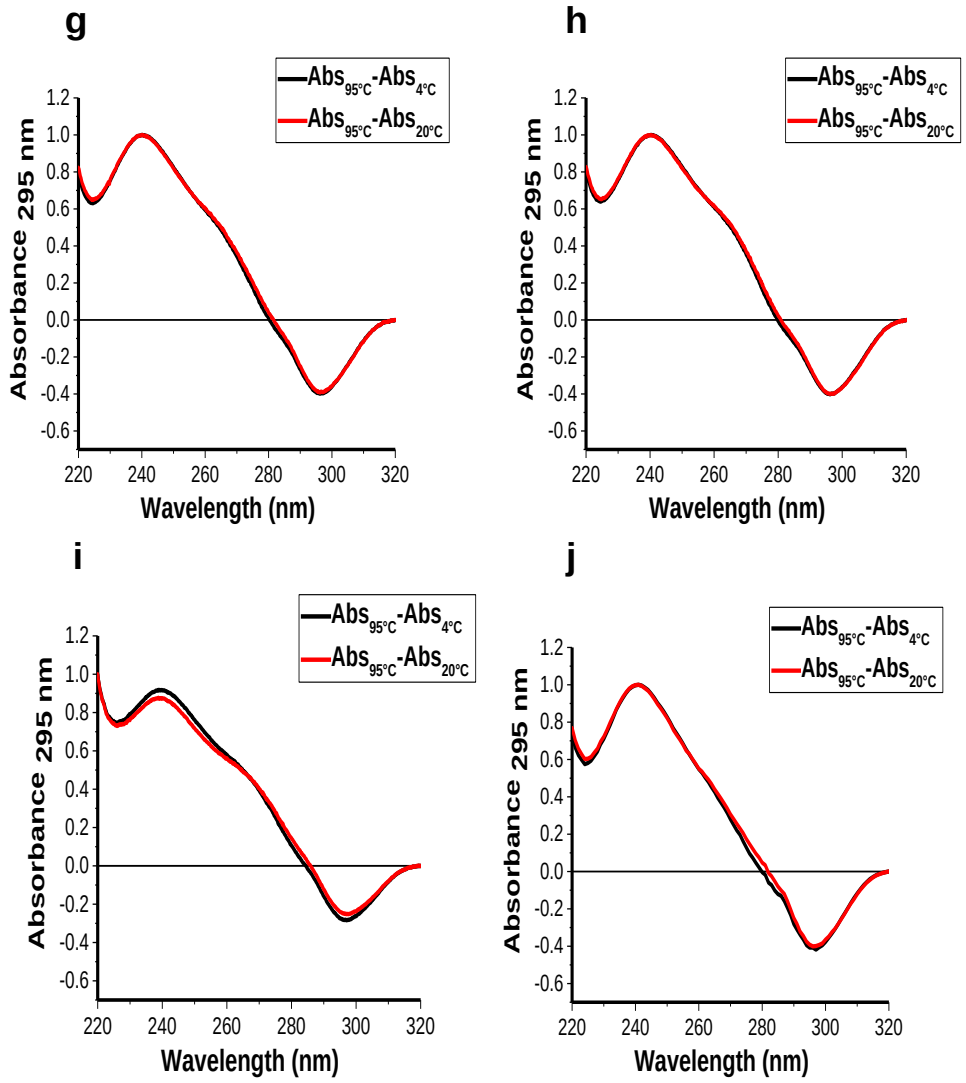


UV melting/annealing profiles of 2.5µM RNA in sodium cacodylate buffer (10 mM pH 5.5) and NaCl (100 mM) (**g**)C_7U_3, (**h**)C_8U_3, (**i**)C_9U_3 and (**j**)$C_{10}U_3$.

A2. The Thermal difference spectra of C_nU_3 at pH 5.5

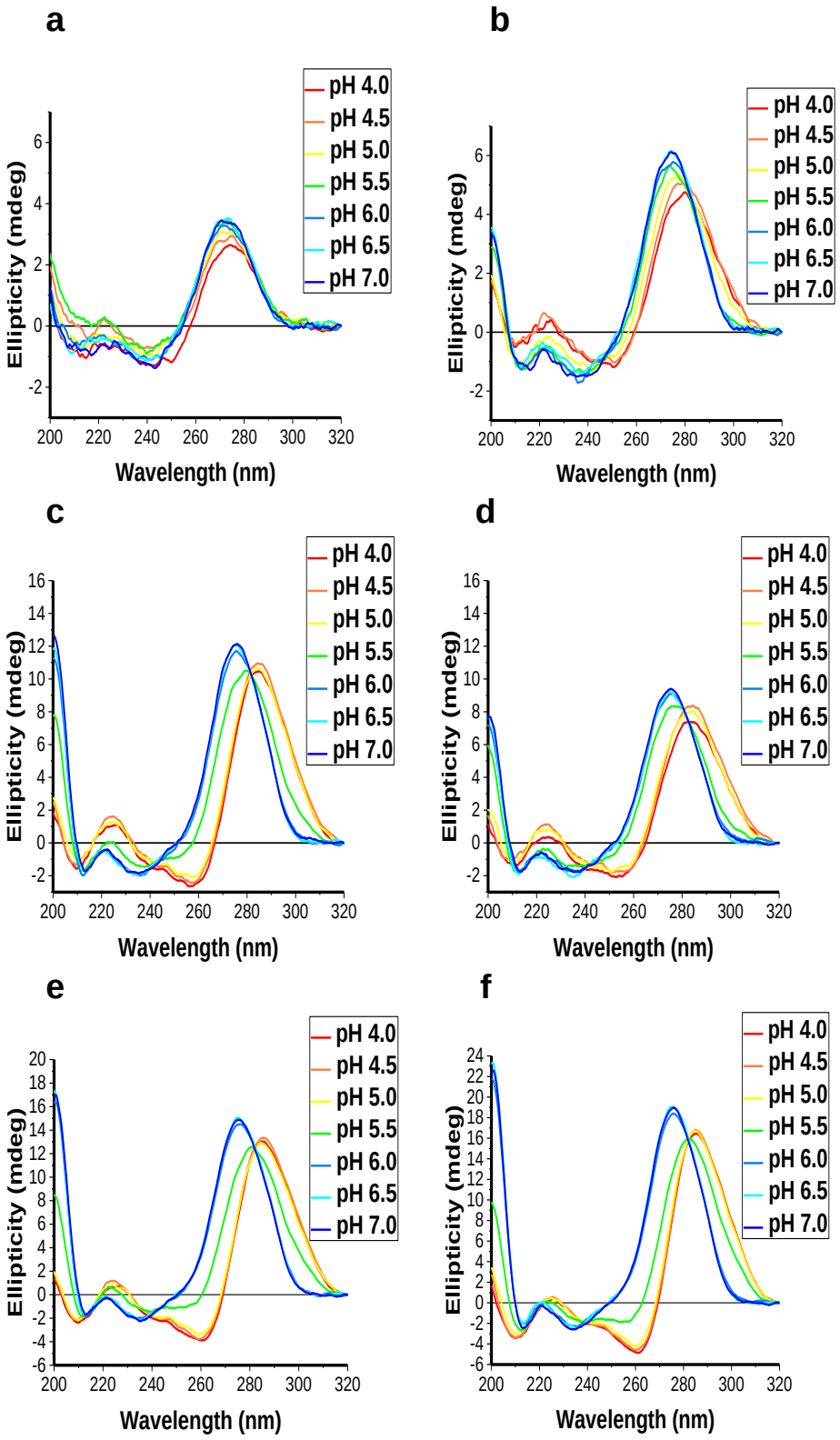


The thermal difference spectra (TDS) profiles of 2.5µM RNA in sodium cacodylate buffer (10 mM pH 5.5) and NaCl (100 mM) of (**a**)C_1U_3, (**b**)C_2U_3, (**c**)C_3U_3, (**d**)C_4U_3, (**e**)C_5U_3, (**f**)C_6U_3.

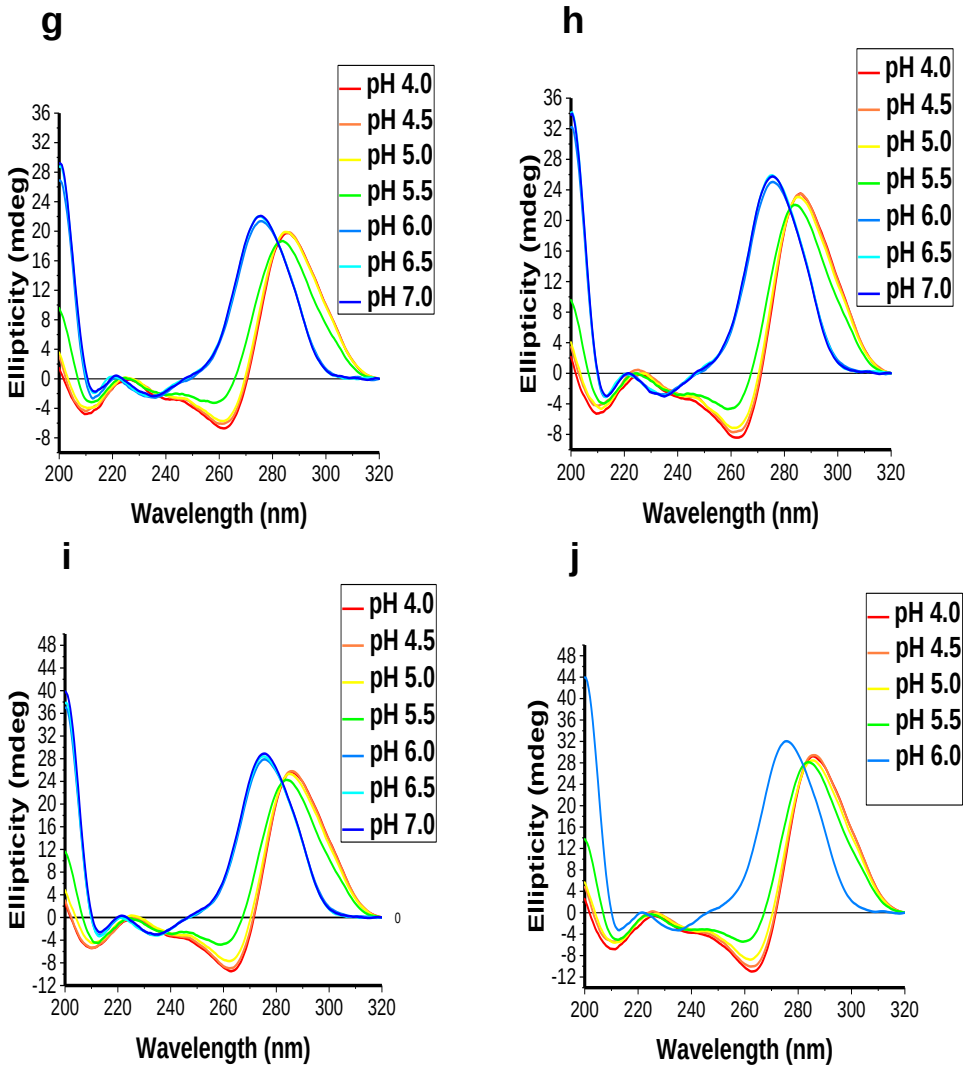


The thermal difference spectra (TDS) calculated between 95 and 4°C profiles of 2.5µM RNA in sodium cacodylate buffer (10 mM pH 5.5) and NaCl (100 mM) of (**g**)C_7U_3, (**h**)C_8U_3, (**i**)C_9U_3 and (**j**)$C_{10}U_3$

A3. The Circular Dichroism of C_nU_3 at pH 5.5

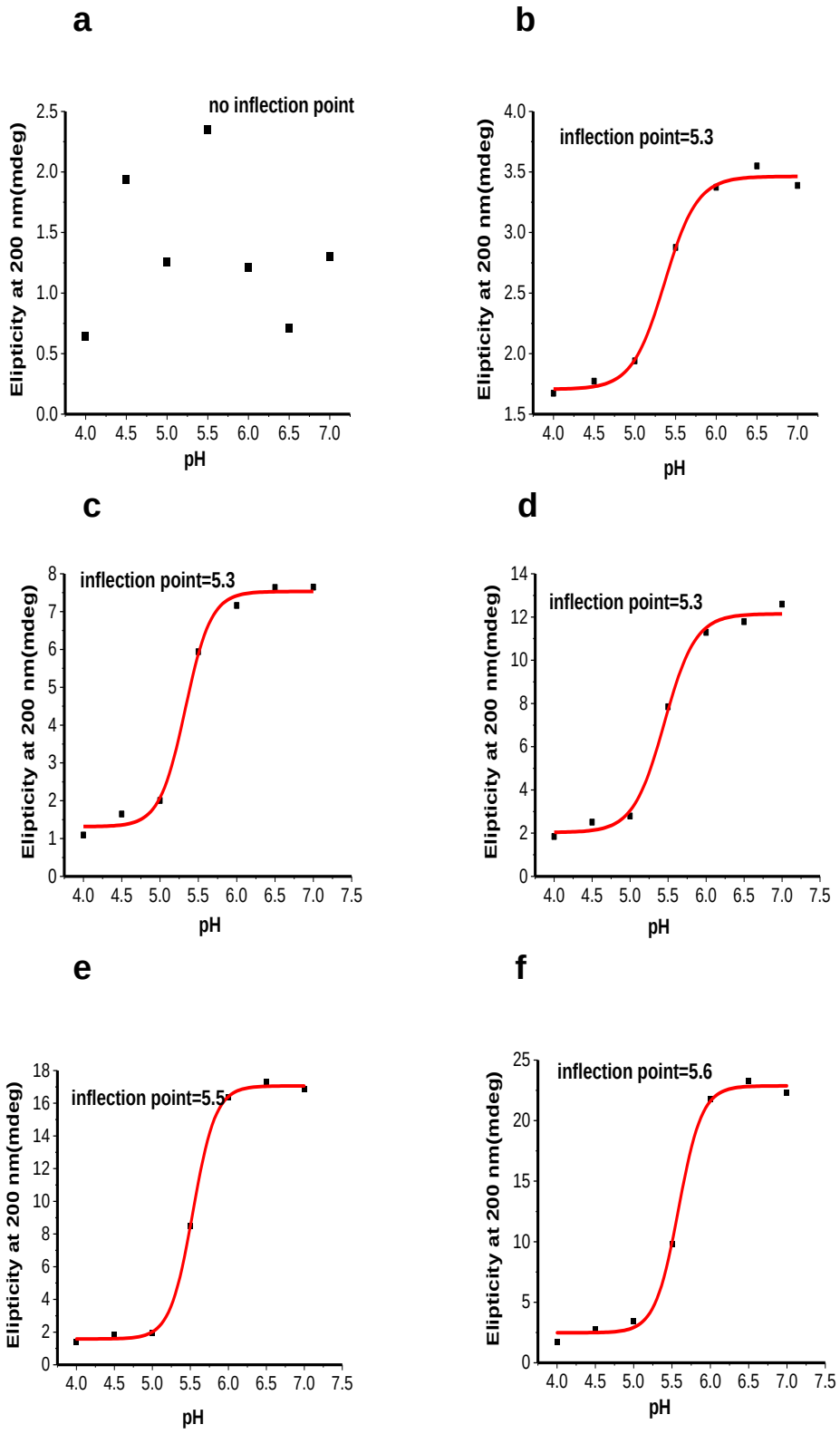


Circular Dichroism of 10 µM RNA in sodium cacodylate buffer (10 mM pH 5.5) and NaCl (100 mM) (**a**)C_1U_3, (**b**)C_2U_3, (**c**)C_3U_3, (**d**)C_4U_3, (**e**)C_5U_3, (**f**)C_6U_3.

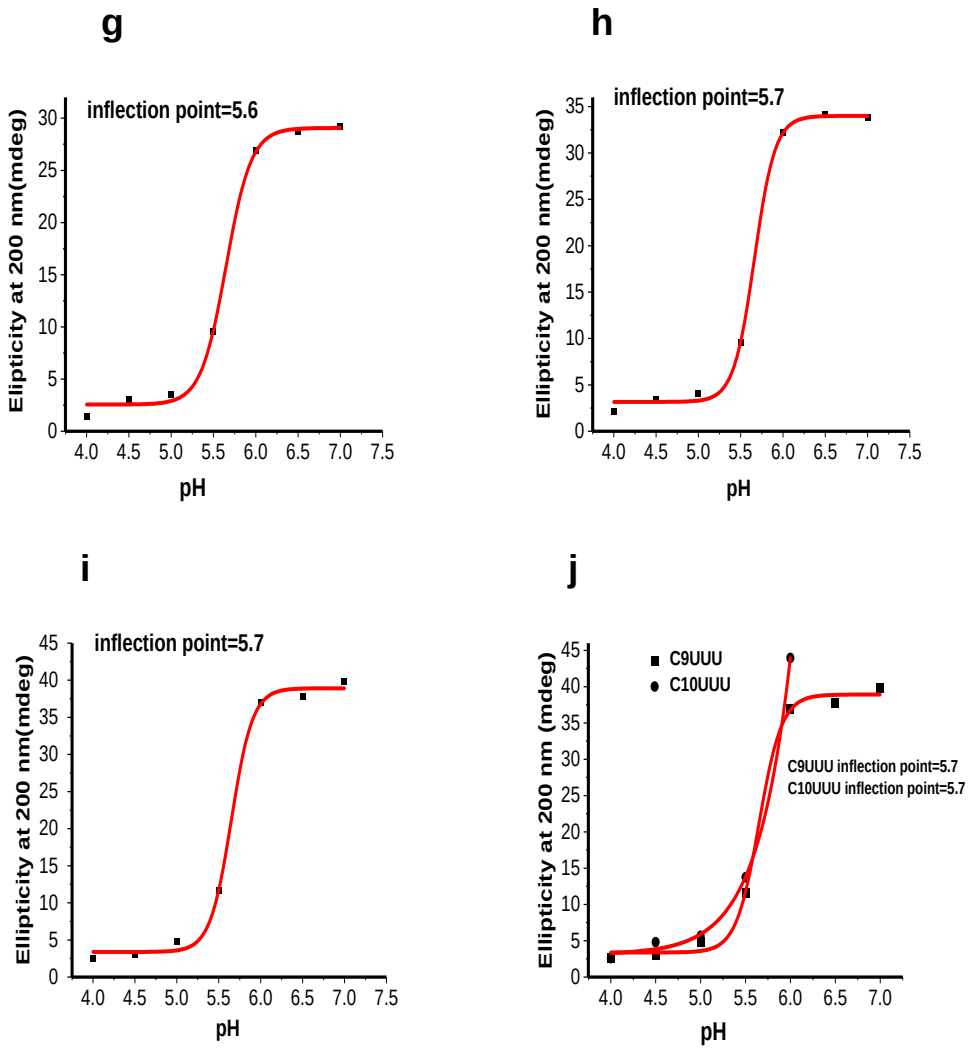


Circular Dichroism of 10 µM RNA in sodium cacodylate buffer (10 mM pH 5.5) and NaCl (100 mM) (**g**)C_7U_3, (**h**)C_8U_3, (**i**)C_9U_3, (**j**)$C_{10}U_3$.

A4. The ellipticity at 200 nm versus pH for all the sequences from our library.

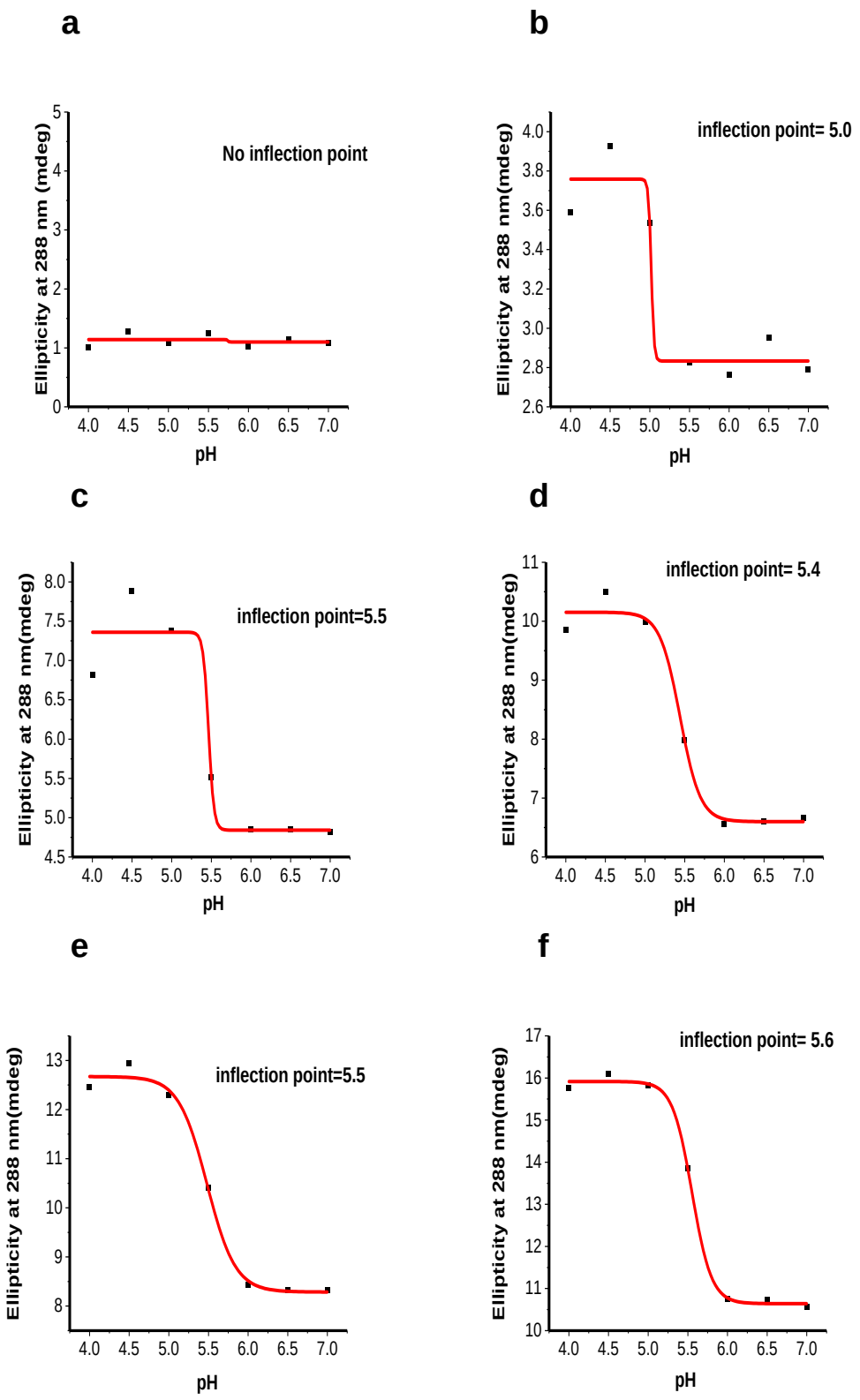


Ellipticity at **200** nm versus pH of 10 µM RNA in sodium cacodylate buffer (10 mM pH 5.5) and NaCl (100 mM), inflection points were calculated by plotting this data in a sigmoidal graph (**a**)C_1U_3, (**b**)C_2U_3, (**c**)C_3U3, (**d**)C_4U_3 (**e**)C_5U_3, (**f**)C_6U_3.

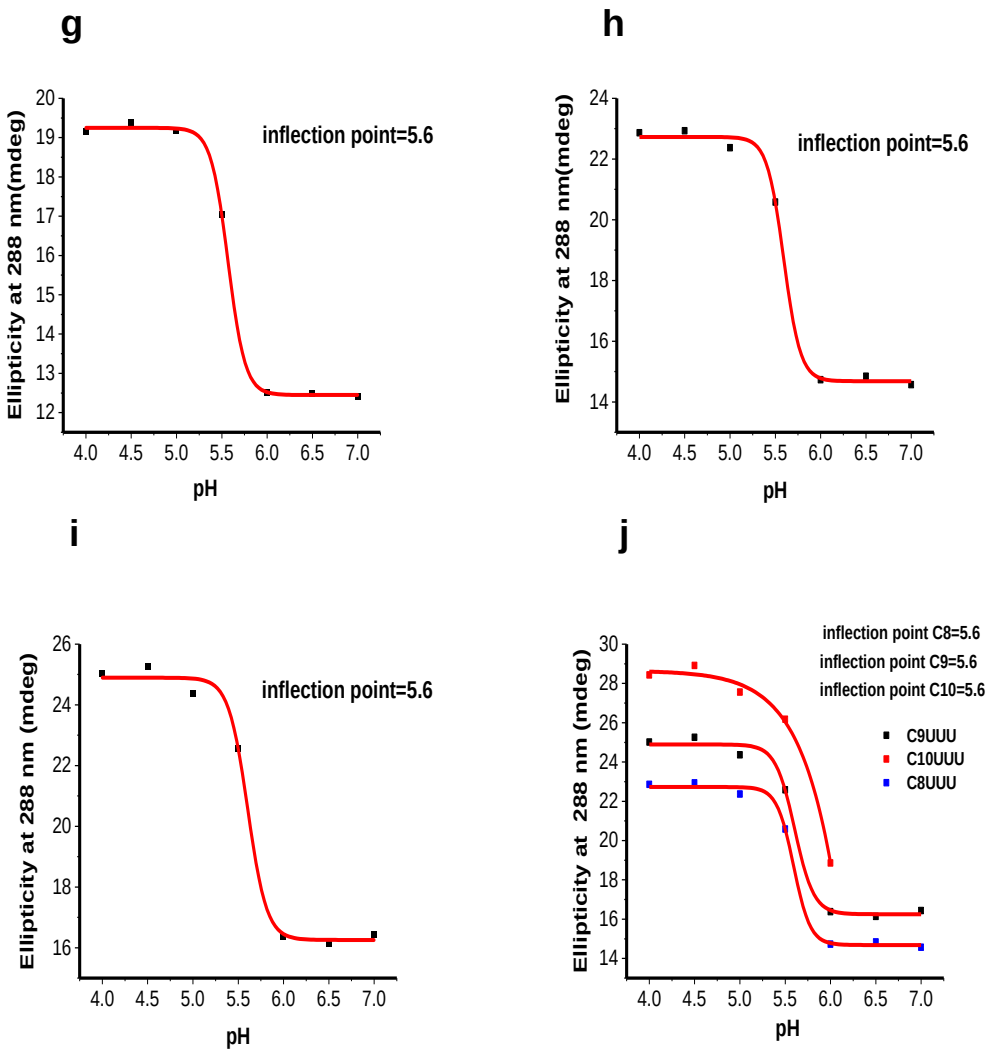


Ellipticity at **200** nm versus pH of 10 µM RNA in sodium cacodylate buffer (10 mM pH 5.5) and NaCl (100 mM), inflection points were calculated by plotting this data in a sigmoidal graph (**g**)C_7U_3, (**h**)C_8U_3 (**i**)C_9U_3, (**j**) C_9U_3 and $C_{10}U_3$.

A5. The ellipticity at 288 nm versus pH for all the sequences from our library.

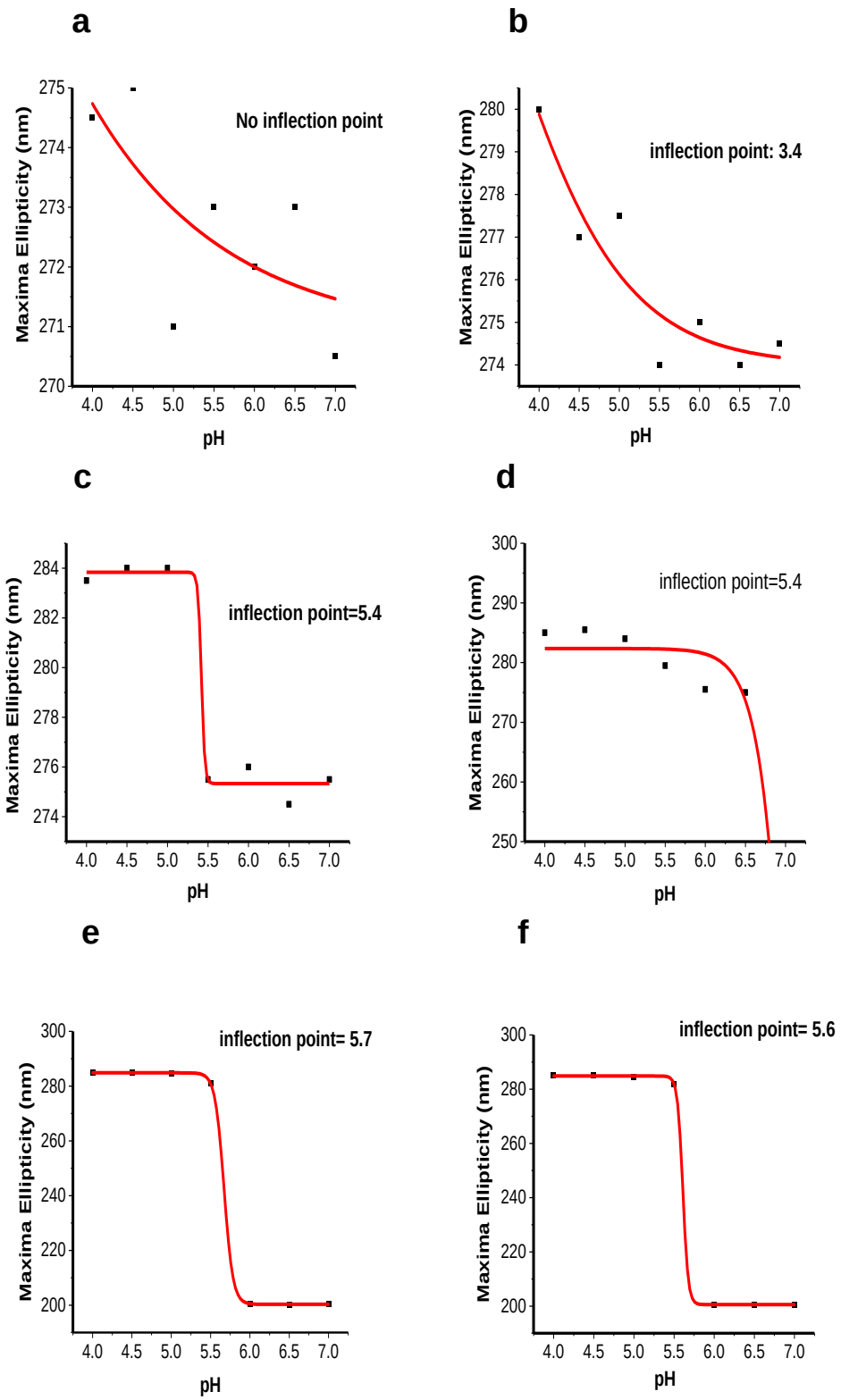


Ellipticity at **288** nm versus pH of 10 µM RNA in sodium cacodylate buffer (10 mM pH 5.5) and NaCl (100 mM) showing the inflection points for (**a**)C_1U_3, (**b**)C_2U_3, (**c**)C_3U3, (**d**)C_4U_3 (**e**)C_5U_3, (**f**)C_6U_3.

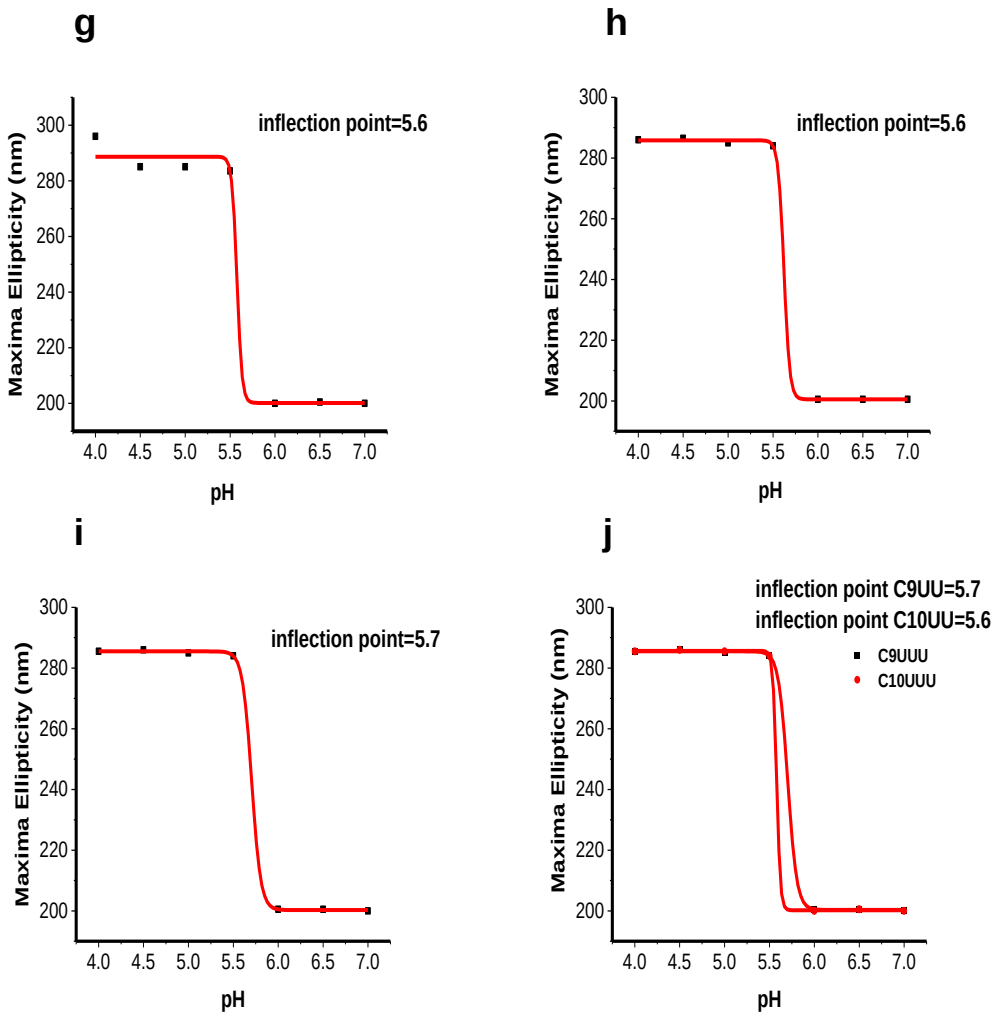


Ellipticity at **288** nm versus pH of 10 µM RNA in sodium cacodylate buffer (10 mM pH 5.5) and NaCl (100 mM), inflection points were calculated by plotting this data in a sigmoidal (**g**)C_7U_3, (**h**)C_8U_3, (**i**)C_9U_3 and (**j**)C_8U_3, C_9U_3 and $C_{10}U_3$.

A6. The maxima ellipticity versus pH for all the sequences from our library.

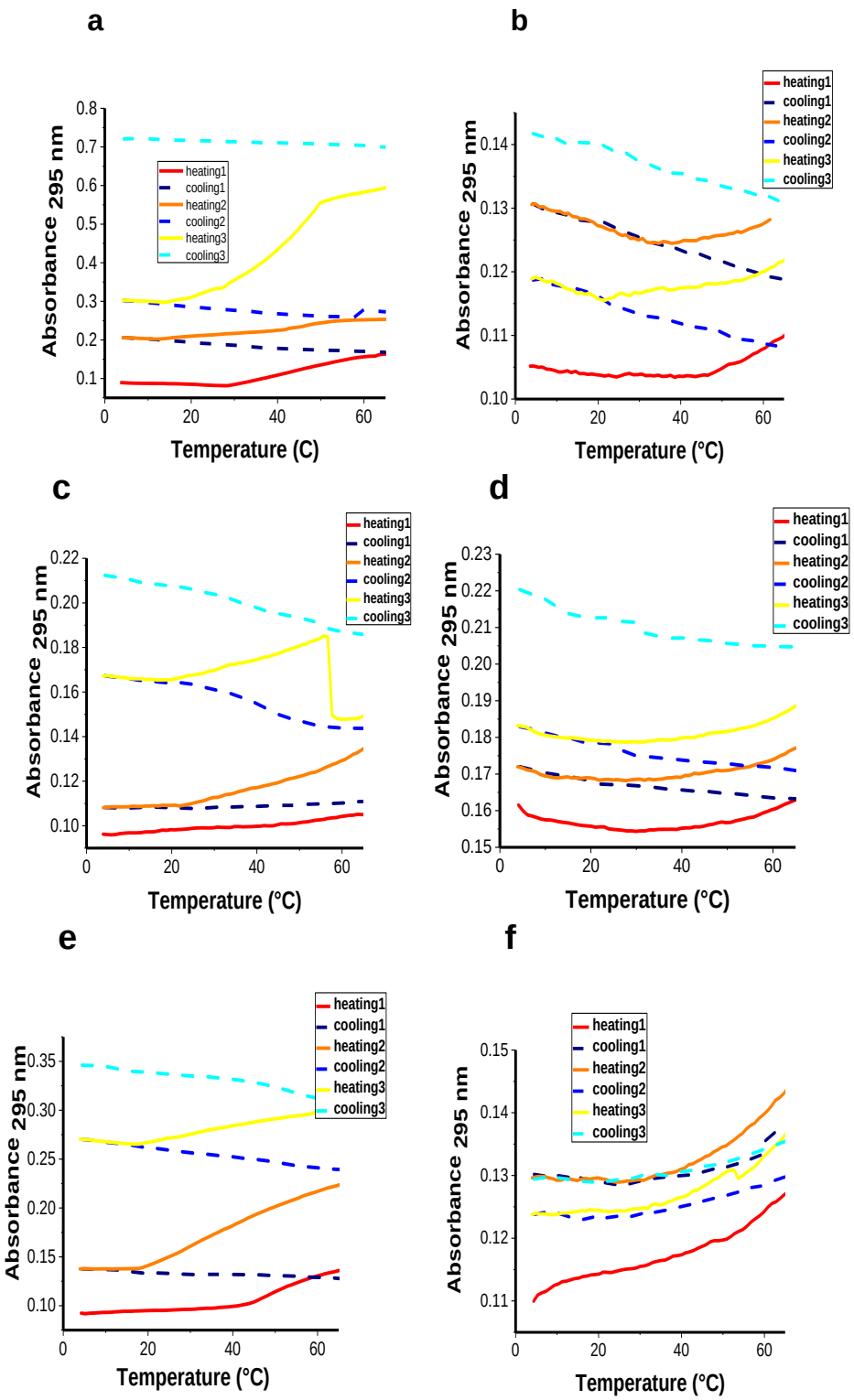


Maxima Ellipticity versus pH of 10 µM RNA in sodium cacodylate buffer (10 mM pH 5.5) and NaCl (100 mM), inflection points were calculated by plotting this data in a sigmoidal graph (**a**)C_1U_3, (**b**)C_2U_3, (**c**)C_3U3, (**d**)C_4U_3, (**e**)C_5U_3, (**f**)C_6U_3.

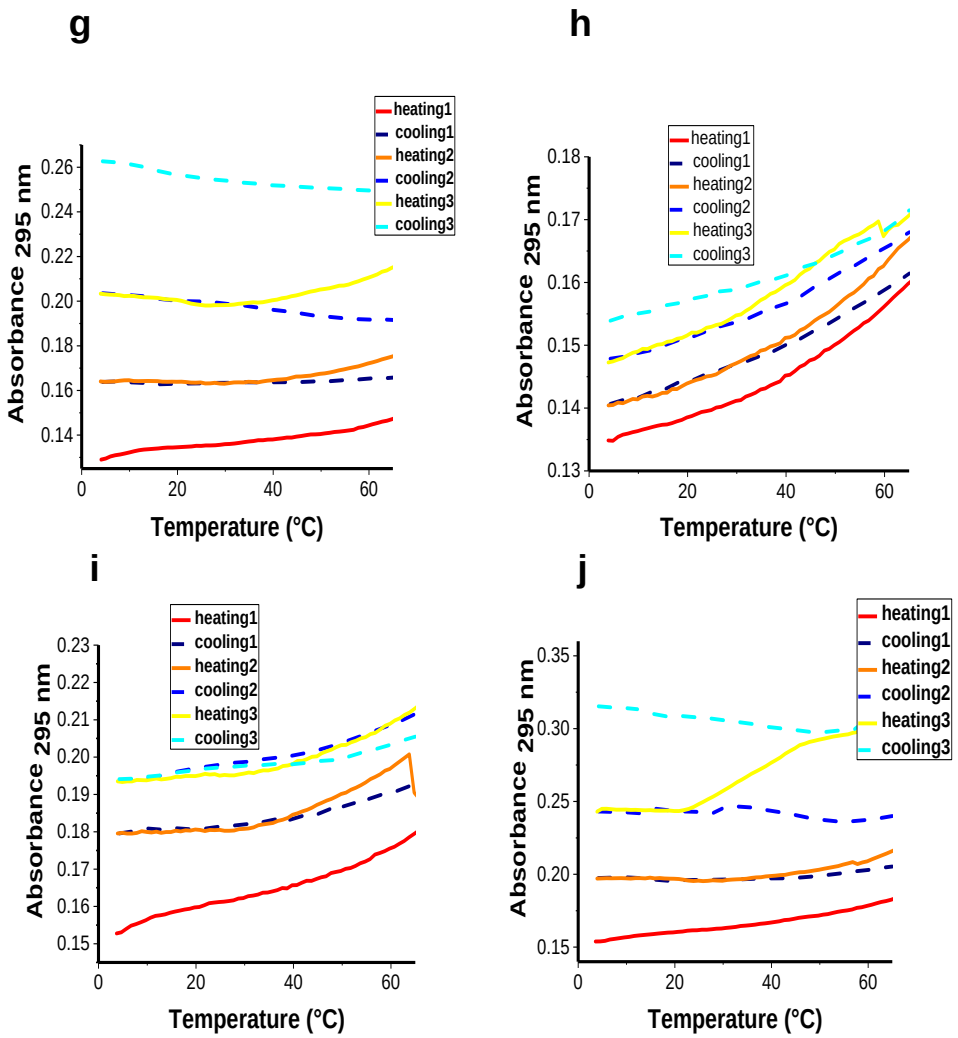


Maxima Ellipticity versus pH of 10 µM RNA in sodium cacodylate buffer (10 mM pH 5.5) and NaCl (100 mM), inflection points were calculated by plotting this data in a sigmoidal graph (**g**)C_7U_3, (**h**)C_8U_3, (**i**)C_9U_3 (**j**) C_9U_3 and $C_{10}U_3$.

A7. UV melting and annealing profile at pH 7.4 for C_nU_3

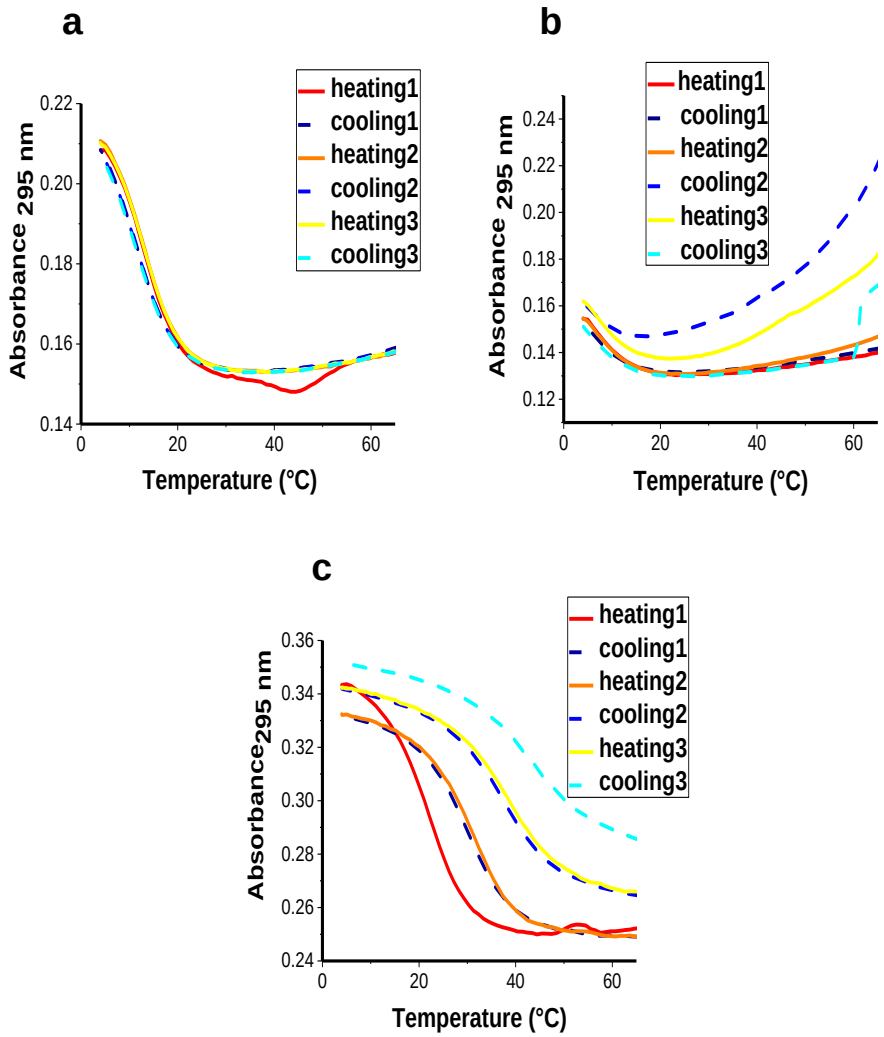


UV melting/annealing profile of 2.5µM RNA in sodium cacodylate buffer (10 mM pH 7.4) and NaCl (100 mM) (**a**)C_1U_3, (**b**)C_2U_3, (**c**)C_3U_3, (**d**)C_4U_3, (**e**)C_5U_3, (**f**)C_6U_3.



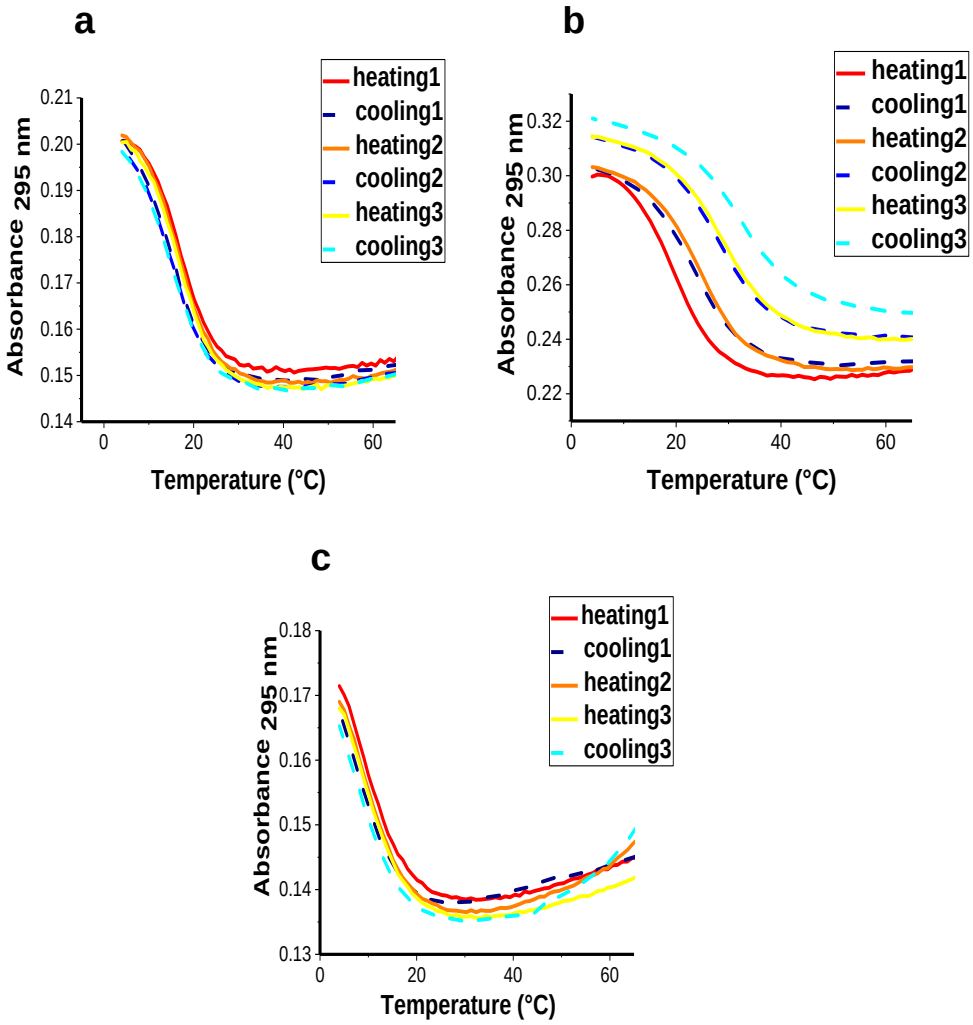
UV melting/annealing profile of 2.5µM RNA in sodium cacodylate buffer (10 mM pH 7.4) and NaCl (100 mM) (**g**)C_7U_3, (**h**)C_8U_3, (**i**)C_9U_3 and (**j**)$C_{10}U_3$.

A8.UV melting/annealing profile AT5G 10 mM buffer pH 5.8 from control experiment.



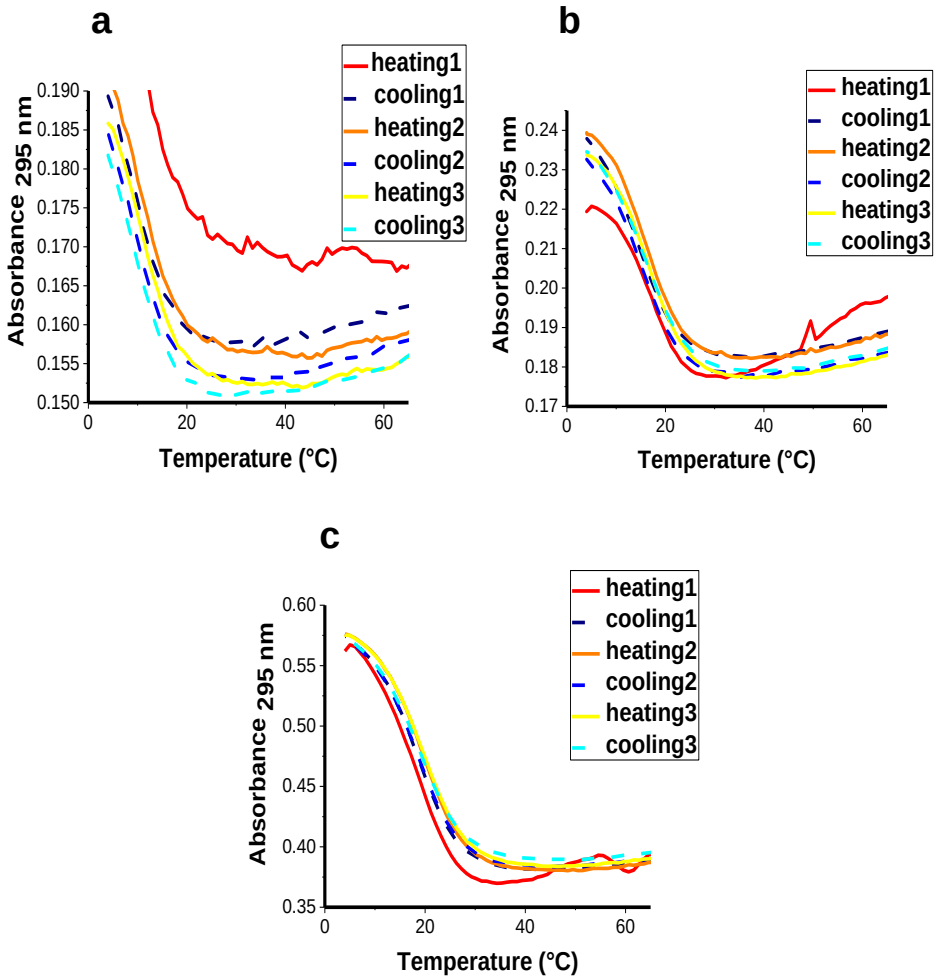
UV melting and annealing profiles of 2.5 µM AT5G RNA and sodium cacodylate buffer (10 mM pH 5.8) (**a**) in KCl (100 mM) (**b**) KCl (800 mM) (**c**) KCl (100 mM) and PEG 8,000(40%).

A9.UV melting/annealing profile AT5G 50 mM buffer pH 5.8.



UV melting and annealing profiles of 2.5 µM AT5G RNA and sodium cacodylate buffer (50 mM pH 5.8) (**a**) in KCl (100 mM) (**b**) KCl (100 mM) and PEG 8,000(20%) (**c**) KCl (800 mM).

A10.UV melting/annealing profile AT5G 50 mM buffer pH 5.8, KCl 100 mM with large stock buffer.



UV melting and annealing profiles of 2.5 µM AT5G RNA and sodium cacodylate buffer (50 mM pH 5.8) (**a**) no PEG (**b**) PEG 8,000(20%) (**c**) PEG 8,000(40%).