Sequence Dependent RNA UNCG and DNA TNCG Tetraloop Hairpin Melting and Refolding Dynamics Using Temperature-Jump Infrared Spectroscopy

Clement Peter Gregory Howe

PhD

University of York

Chemistry

January 2023

Abstract

Temperature-jump Infrared spectroscopy (T-jump IR) is a label-free method that provides access to non-equilibrium dynamic processes on timescales of nanoseconds to milliseconds, capturing changes in vibrational modes that indicate changes in molecular structure.

This thesis outlines the development of T-jump IR methodologies to study RNA UNCG tetraloop hairpins, an RNA motif important for proper folding of its tertiary structure and stability. To gain insight into the structural basis for their relative behaviour, analogous DNA hairpins were also studied. Chapter 3 presents a T-jump methodology to measure RNA and DNA melting and refolding dynamics in a single measurement for the first time. This established that RNA and DNA hairpins melt on different timescales. Their refolding times however were similar, with the formation of base-stacks the rate limiting step for both.

In Chapter 4, the influences of the stem sequence on the RNA and DNA hairpin dynamics are investigated by introducing an AT pair into the stem sequence to provide site-specific information. By tracking the dynamic changes at particular positions in the stem, the pathways of melting and refolding can be established by comparison. These experiments determined the RNA and DNA tetraloops melt and refold via different intermediate conformations.

Chapter 5 extends the study of the UNCG tetraloop dynamics using Two-dimensional infrared spectroscopy (2D-IR) and T-jump 2D-IR in order to observe the changes in the relationships of the vibrational modes during the melting dynamics. By doing this a greater level of structural detail can be obtained, allowing more precise determination of the conformational dynamics over the melting process. This work suggested DNA melting follows a more multistate transition than RNA, due to greater flexibility in the stem.

Table of Contents

Acknowledgements

I would like to thank my supervisor Neil Hunt for his guidance through all the work I've done over the last four years. He has always been open to talk out any problems I encountered, challenged me to be better, and known when to let me muddle through something on my own. Through lockdowns and working from home, he has always been quick to respond or set up a video chat when necessary. I would also like to thank Tony , for always asking the questions that stump me, making me really think and for teaching me that any problem can be solved with an excel spreadsheet.

To the members of our research group, Barby, Sol, Soheila and Clara it's been a genuine pleasure. We had some fun times and some tough times together and many long nights doing experiments. Barby in particular has been there for the whole journey and has been a constant source of support and encouragement throughout.

I would like to express my gratitude to the team down at RAL for their support; Greg, Mike, Emma, Marta, Ian and Paul, without their knowledge and expertise of the experiments and ability to find solutions and workarounds to problems I encountered, none of this would have been possible.

I'd like to recognise the inspirational teachers I had along the way, Dr. Michael Friedeberg, Dr. Ian Gameson and Dr. Jason Camara.

Finally I would like to give a special thank you to my mother Elizabeth and partner Mansi for their support and patience particularly over my writing up period.

Declaration

I declare that this thesis is a presentation of original work and I am the sole author. This work has not previously been presented for an award at this, or any other, University. All sources are acknowledged as References.

Chapter 3 contains results published in the following publication:

C. P. Howe, G. M. Greetham, B. Procacci, A. W. Parker and N. T. Hunt, J. Phys. Chem. Lett., 2022, **13**, 9171–9176.

Chapter 4 contains results published in the following publication:

C. P. Howe, G. M. Greetham, B. Procacci, A. W. Parker and N. T. Hunt, J. Phys. Chem. B, 2023, **127**, 1586–1597.

1. Introduction and Theory

The oligonucleotide deoxyribonucleic acid (DNA) is fundamental to living cells for the storage and coding of genetic information of organisms in the sequence of bases which make up its structure.[1,2] In particular it contains the blueprints for proteins. In the central dogma of molecular biology, this genetic information is transferred in two ways; DNA replicates to form more copies of itself to preserve genetic information or it is transcribed into a ribonucleic acid (RNA) sequence (Fig. 1.1).[1] The RNA base sequence can then be translated into an amino acid sequence from which proteins are assembled or form a functional RNA structure. [1,2]

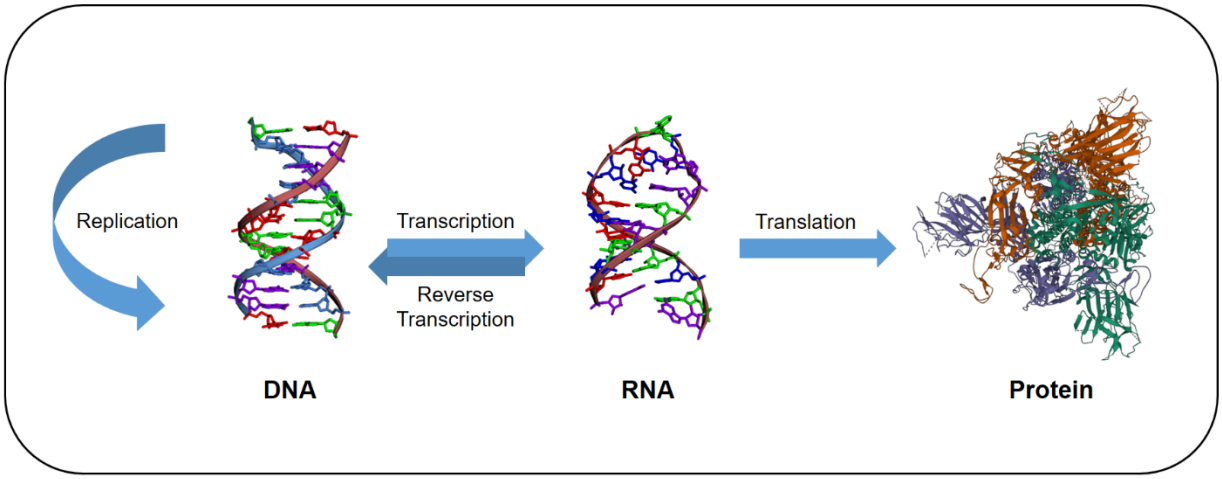

Figure 1.1: *The Central Dogma of Molecular biology. DNA: 1ILC,*[3] *RNA: 1MFJ,*[4] *Protein: 6XLU.*[5] *Structures visualised using EzMol.*

These processes are highly dynamic, with the replication of eukaryotic DNA progressing at a rate of 50 nucleotides per second. In its double stranded form, this information in DNA is actually inert, but helicase enzymes “unzip” the strands of DNA allowing access to the bases.[6] Polymerase enzymes then bind to these open sections in order to replicate DNA or in the transcription of the bases into RNA.[7] Disruption to these critical functions can cause carcinogenesis or

genetic disorders, so understanding how these processes work, and how they are repaired and maintained in biological systems is fundamental.[8]

The role of RNA is just as critical, performing a range of catalytic functions as well as being capable of ligating or cleaving the bonds of other biomolecules in addition to its role in protein synthesis.[9] By targeted intervention in its biological processes it's possible to manipulate RNA function to promote health and fight disease. For instance RNA instructions can be introduced for new proteins, such as in the case of mRNA based vaccines where these proteins are used to train the immune systems to fight viral infection.[10,11] Conversely, antisense therapies can be used to target RNAs in order to stop the production of specific proteins, such as mutants that are responsible for conditions such as muscular dystrophy,[12] and RNA binding proteins (RBP) can be used to regulate the behaviour of RNA and its functions, the dysregulation of which is associated with a variety of cancers.[13]

These two nucleic acids are active and dynamic components of biological systems. The work presented in this thesis aims to provide insight into the dynamics behaviour of DNA and RNA, thus furthering our understanding of the biological processes at the heart of their functional roles, opening up the possibility of refining existing techniques as well as developing new, targeted medical interventions and therapies.

1.1 Nucleic Acids

To study the behaviour of DNA and RNA, it is first necessary to understand the foundations of their molecular structure. Nucleic acids are biological polymers whose primary structure consists of one of five nucleotide monomers. Nucleotides are made up of a ribose sugar derivative, with a nitrogenous

nucleobase bound to the 1’ position and a phosphate unit to the 5’ position (Fig. 1.2). The phosphate groups are bound to the 3’ position of the next nucleotide monomer. These form as long chains, or strands, with defined sequences of nucleotides.

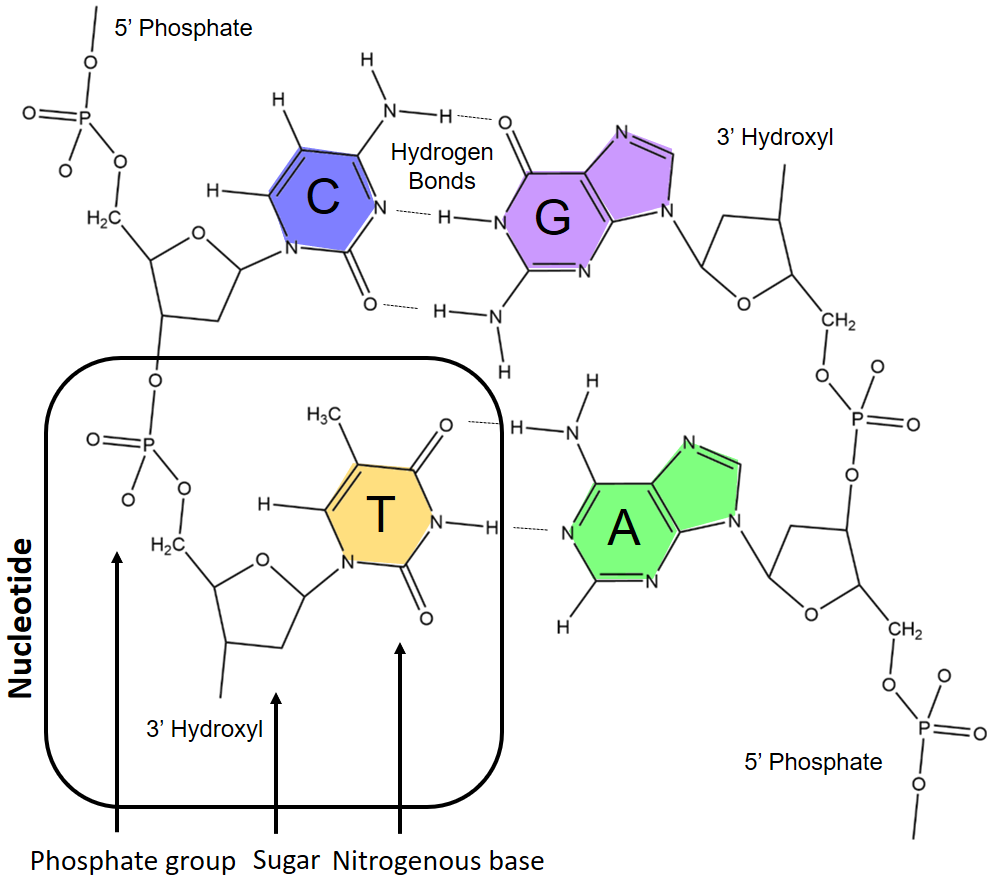

Figure 1.2: *Diagram of the structure of two complementary, antiparallel base-paired DNA strands. A nucleotide sub-unit has been highlighted to show the constituent nitrogenous base, sugar and phosphate groups. Hydrogen bonding can be seen between guanine (purple) and cytosine (blue), and adenine (green) and thymine (orange) forming Watson-Crick base pairs.*

The nucleobases bound to the sugar units come in two types, the pyrimidines, cytosine and thymine or uracil, and the purines, guanine and adenine. Due to the ability of the bases for form hydrogen bonds, the bases form pairs consisting of a purine and pyrimidine, specifically adenine pairs with thymine (DNA) or

uracil (RNA), forming two hydrogen bonds between them, and guanine and cytosine pair with three hydrogen bonds. These are called Watson-Crick (W-C) base pairs.[14]

This pairing is responsible for the secondary structure of nucleic acids, in which two separate strands with complementary base pair sequence can hydrogen bond to one another in an antiparallel configuration (Fig. 1.2), where the base at the 5' end of a strand forms a W-C pair with the matching base on the 3' end of the second strand. In the case of DNA, the crystal structure was first resolved using x-ray diffraction, determining atomistic detail of the structure and famously deriving the double helical conformation that the double stranded structure adopts.[14]

In the helical conformation (Fig. 1.3), base-stacking between bases provides the primary source of stability to the structure,[15] with the off-set π-stacking of the DNA bases on the order of 3-14 kJ mol^{-1}.[16,17] The structure is further stabilised by base-pairing, which in both RNA and DNA contributes 5-10 kJ mol^{-1}.[1] These are only slightly above the 2.6 kJ mol^{-1} available under room temperature conditions. This provides sufficient stability, while maintaining a minimal barrier to critical interactions, such as enabling ease of "unzipping" of the double helix for replication and transcription .[6]

Under physiological conditions, DNA typically adopts so called B-form helices (Fig. 1.3). This conformation is typified by having 10 base per helical turn and an inter-base distance of 3.3 Å. By contrast, A-form helices are preferentially formed in RNA, and are less tightly coiled but with tighter base stacking with 11 bases per turn and an inter-base distance 2.3 Å (Fig. 1.3). This conformation is less common in DNA, but may be biologically important in such functions as transcription.[18]

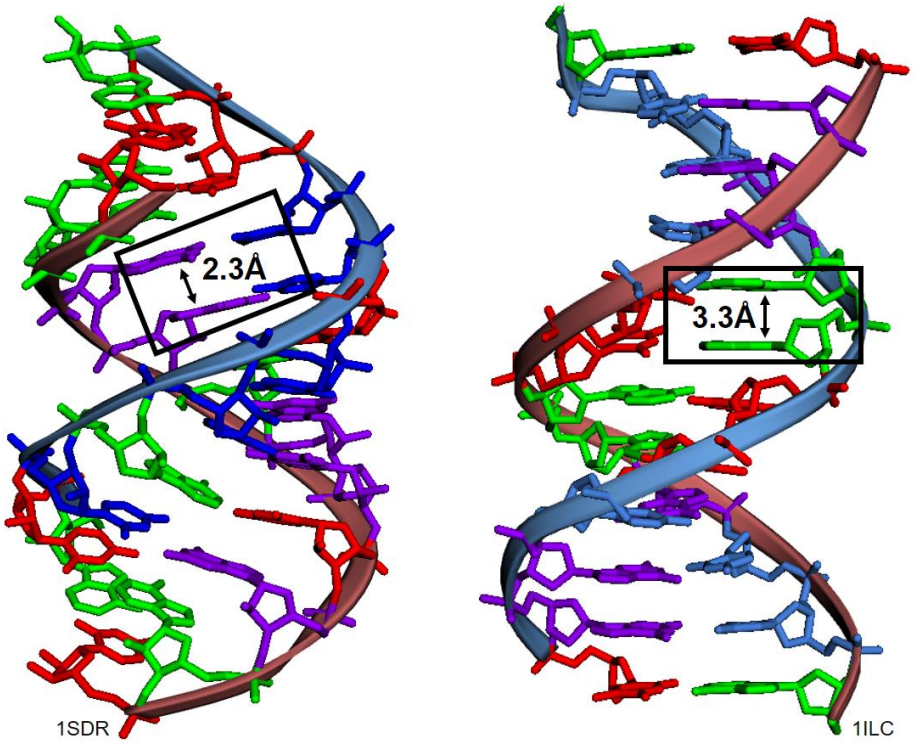

Figure 1.3: *Native helical conformations of double stranded oligonucleotides in solution; RNA typically adopts A-form helices, while DNA prefers B-form. Structures shown are, A-form: 1SDR,*[19] *B-form: 1ILC.*[3] *Boxed region highlights the base stacking in the different conformations.*

1.1.1 Comparing RNA and DNA structure

Although RNA and DNA are very similar in many respects, when it comes to characterising them and particularly comparing them, there are slight differences; DNA uses the 2'-deoxy ribose sugar in its backbone structure, RNA uses ribose (Fig. 1.4 (a)) and DNA utilises the pyrimidine base thymine while RNA uses uracil in its place (Fig. 1.4 (b)), which differs compared to thymine by a methyl group.

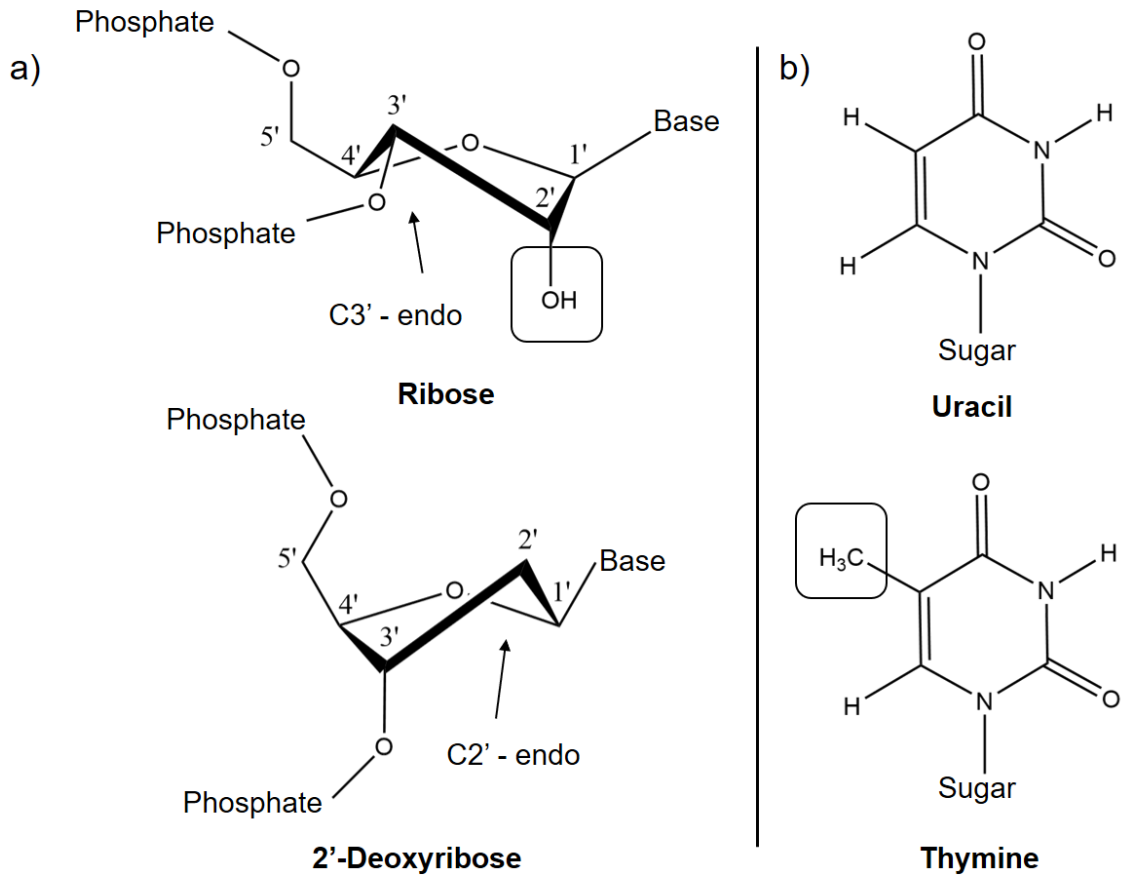

Figure 1.4: *Differences in DNA and RNA structure, a) RNA Ribose and DNA 2'-Deoxyribose sugar conformations, and b) RNA Uracil and DNA Thymine bases. Boxed areas highlight the structural differences, the 2'-hydroxyl of the ribose sugar and the methyl group of the thymine base.*

These difference may seem slight, but it results in some fundamental physical differences. DNA deoxyribose sugars in solution adopt 2'-endo puckering conformations which are responsible for the formations of B-form DNA double helical structure, whereas RNA ribose adopts the 3'-endo conformation, which influences the packing of the phosphate groups and results in preferential formation of A-form helices (Fig. 1.4 (a)).[18,20]

Although RNA does form double stranded helices, it typically exists in single strands folded into functional three dimensional, or tertiary, structures. This

more complex biological behaviour allows RNA to take a variety of forms that fall into three classes: messenger RNA (mRNA), ribosomal RNA (rRNA) and transfer RNA (tRNA).[2] mRNA contains a copy of a section of DNA that encodes for a protein, tRNA match the base code of mRNA to a sequence of amino acids, while rRNA in the ribosomes facilitates the synthesis of proteins from the amino acids. In its different forms, some secondary structural elements are more widely represented across the different RNAs, including such motifs as bulges, internal loops, junctions, single stranded regions and myriad hairpin structures (Fig. 1.5).[21–29]

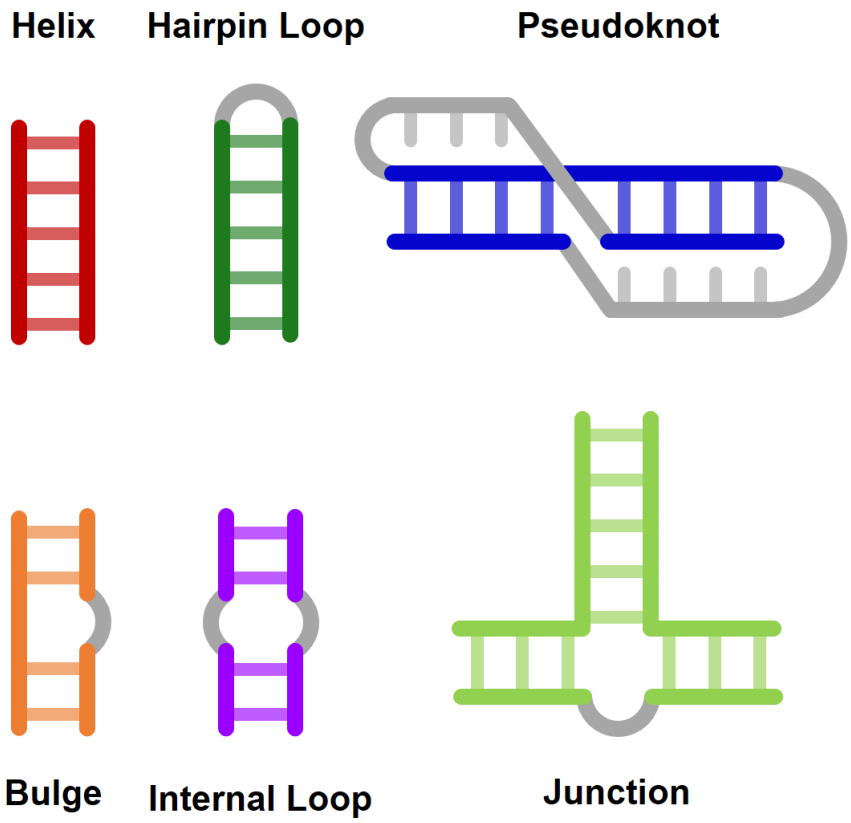

Figure 1.5: *Common RNA secondary structural motifs, with diverse functional structures generated from different combinations of single and double stranded domains.* [21,29]

1.1.2 The RNA UNCG Tetraloop Motif

The work presented in this thesis will focus on the UNCG tetraloop hairpin structure. Hairpins, comprised of a region of single stranded bases closed by a Watson-Crick base paired stem, are a particularly common, functional feature found in RNA that are important as nucleation sites in the formation of the tertiary structure and as recognitions site for RNA binding proteins .[22,30] While the number of unpaired bases in hairpins can vary, the most common is four, making up more than 55% of all loops in rRNA.[22,31] These are called tetraloops, the vast majority of which conform to three generic loop types, GNRA, UNCG and more rarely CUUG (N = any nucleotide, R = purine). Although GNRA tetraloop motifs are more common in RNA, the UNCG loop is significant due to its unusual thermodynamic stability. The C(UUCG)G loop for instance has a melting point 20 ^{o}C higher than a loop with "normal" stability, such as C(UUUU)G.[32] The structure of the loop, which differs from most loops, may be important in explaining this (Fig. 1.6).

Extensive base-stacking is seen throughout, with the cytosine of the closing pair stacking with the first and third residue and the loop closing guanine stacking with the fourth residue (Fig. 1.6 (b)). The loop is further stabilised by the UG base pairing of the first and fourth loop residue, where the guanine is in the syn conformation, and by an O4'-π contact between the third and fourth residues.[33,34] The UG "wobble" base pair is found in all classes of RNA and has comparable stability to a canonical W-C base-pair (Fig. 1.6 (c)).[35,36] The combination of these interactions results in a so called "Z-turn". Due to this, the loop cannot easily twist, bend or deform, it is inflexible, with the second and third nucleotides primarily fixed in pure 2'-endo sugar conformations. The second and third nucleotides in GNRA loops by contrast show more flexibility with evenly distributed 3'-endo and 2'-endo conformations.

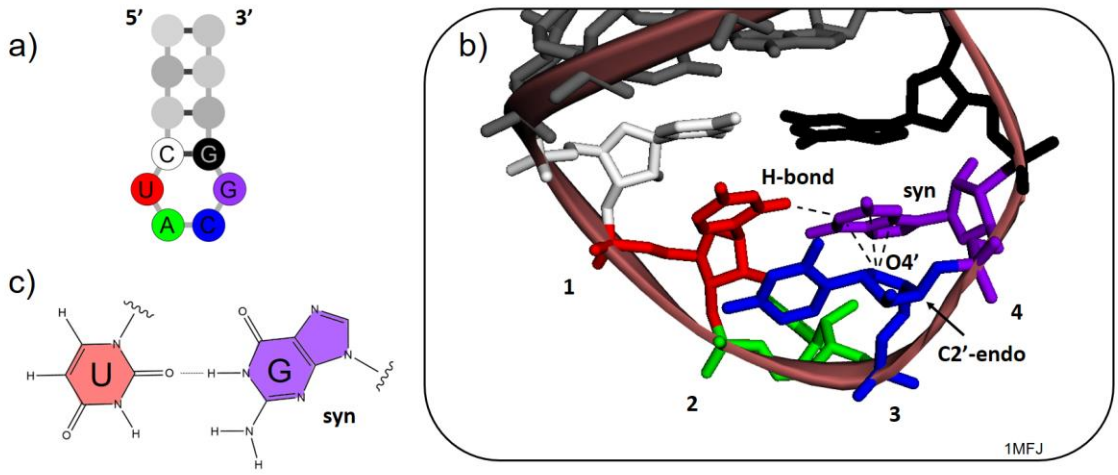

Figure 1.6: *Diagram showing a) a schematic representation of a UNCG hairpin, b) structure of the UNCG Z-turn, including an embedded trans-Sugar (t-SW) interaction between the first (red) and fourth (purple) nucleobase forming a UG wobble pair, facilitated by the C2'-endo puckering of the third residue (blue), and the third and fourth (blue and purple) ribose rings are configured in a head-to-tail orientation resulting in a O4'- π stacking contact, and c) the configuration of the UG wobble pair.*

One of the more studied UNCG sequences is 5'-GGAC(UUCG)GUCC-3', which has been the subject of a series of publications by Tinoco, Jr. et al.[33,37,38] In their work they demonstrated that a TTCG DNA loop sequence had a melting point 18 ^{o}C lower than the analogous RNA sequence. Thermodynamic studies using RNA with dU in the loop determined this discrepancy in thermal stability is due to the central two ribose sugars in the loop being fixed in the 2'-endo conformation, while deoxyribose in the loop is able to easily form either the C3'-endo or C2'-endo conformations. This suggests the 2'-hydroxyls are key in the loop's stability (Figs 1.4 (a), 1.6 (b)).[23]

1.1.3 Nucleic Acid Analysis

The molecular composition, primary and secondary structure of nucleic acids lend themselves to investigation by a wide variety of different methods. Along with X-ray diffraction, NMR is also capable of resolving atomistic structural detail of nucleic acids, with the added benefit of being able to study them in the solution phase. NMR studies of DNA and RNA structures have provided insight into the solution phase conformations, and can determine fine detail such as the base-pairing interactions in nucleic acids, and determine secondary structure and surface interactions with ligands and proteins.[39,40] Multi-dimensional NMR in particular has been crucial for solving complex oligonucleotide structures in solution with the use of isotopic labelling. [30,41–45] Using NMR it has been possible to resolve specialised hairpin loop structures used in protein recognition in RNA.

Base-stacking in the helical structure produces a characteristic response in the UV range, with a sharp rise in absorbance at 260 nm observed on the denaturation of the DNA strands, called the hyperchromic effect. This is caused by changes in the localization of the electronic excited states between the single-stranded and base-paired conformations.[46–49] This effect has been integral in the studying the thermodynamics of DNA and RNA with UV-Vis spectroscopy.[9,38,50,51] With the introduction of fluorescent labelled bases, site specific information has also been gathered on secondary and tertiary structure.[52,53]

Due to the asymmetry of the backbone sugars and their formation of right handed helices, DNA and RNA are optically active.[54] This makes them open to investigation using optical rotatory dispersion and circular dichroism. The specific base sequence and conformation cause differing degrees of rotation of polarised light in the 180 – 320 nm region. [9,45,54–56] Circular dichroism has been

particularly useful for monitoring structural changes as a results of different environmental conditions such as temperature, pH and ion concentration.

Vibrational spectroscopy has been particularly successful at providing label-free structural detail in the solution phase. While it does not provide atomistic detail, it is sensitive to the functional components present in nucleic acids; the sugar-phosphate backbone and each of the nucleotide bases display a variety of distinct Infrared and Raman active modes. [57–61] This make it possible to observe the sugar puckering conformations, the conformations of the phosphate backbone and the individual state of each type of base, its base-pairing and base-stacking in a single measurement and their distinct relationship with their environment.[25,58–60,62,63] Infrared methods have been shown to be highly versatile in the study of nucleic acids, with measurement possible in solutions, films, crystals and other media, as well as being suitable to analyse short sequences as readily as high molecular weight oligonucleotides.[63–65]

In addition, the vibrational modes also couple, providing spatial information about the relationships of the functional groups to each other which can be readily observed with the use 2D-IR methods,[59–61,66–73] making it possible to characterise the relationship of solvent hydrogen bonds to the backbone phosphate groups and the influence of ions on this interaction,[73,74] as well as energy transfer pathways between the bases and the backbone,[61] and the difference in inter and intra-strand coupling of the base modes in double stranded conformations.[70]

Nucleic Acid Dynamics

The role of DNA and RNA function relies not just on their structure, but on their dynamic conformations that interact with, and respond to, changes in their environment: formation and loss of secondary structure and conformational changes in those structures. To study these processes, methods are required that can measure the conformational changes in nucleic acids on the timescales of these processes. Studies of DNA and RNA dynamics have shown that disassociation and association processes of double stranded structures typically occur on timescale of nanoseconds to milliseconds.[52,75–83]

Time resolved NMR has allowed some insight into the conformational dynamics and other dynamic processes of nucleic acids.[84–87] In particular it can probe slow conformational changes, such as the mRNA guanine-sensing riboswitch aptamer domain conformation response to binding hypoxanthine which undergoes a series of folding steps on the timescale of seconds.[84] The time resolution in NMR however is limited to milliseconds to minutes and longer and is not suitable for characterising the faster processes common to nucleic acids.

One method for interrogating these process is by using spectroscopies with nanosecond or greater time resolution in conjunction with a perturbation of the system, typically by introducing a temperature-jump (T-jump). A T-jump consists of inducing a fast rise in the temperature of the solvent, on the scale of nanoseconds or faster. This rapid change in the temperature of the environment causes the nucleic acids in solution to respond to the change in conditions as the thermodynamic barriers maintaining the conformation are overcome, initiating conformational changes such as strand dissociation, or “melting”. This process can then be probed in real time as the process progresses, giving insight into the dynamic properties of the system, the pathways of conformational change, the rates and rate limiting steps.

UV Fluorescence spectroscopies have been widely employed in T-jump methods. It provides time resolution fast enough to be able to capture nanosecond nucleic acid conformational dynamics and the ability to use fluorescent analogues of the nucleic acid bases to provide site specific dynamics. This has been used to explore the melting of DNA and RNA sequences, determining that it is not a two state process, where any point in the transition can be described by the proportion of species in the double or single stranded conformations. Instead it was found to involve multiple intermediate conformations.[88] Some studies have classified distinct end-fraying and unstacking phases of the melting process before full unfolding.[89]

A number of UV studies in RNA in particular have focused on the dynamic properties of different RNA loop structures on melting,[52,81,90–93] with loop sequence and size found to influence the intra-loop interaction of the bases and therefore their ability to stabilise the structure and impede denaturation.

T-jump IR has also proven particularly effective, providing dynamic insight into these non-equilibrium dynamic processes from the perspective of the different bases directly involved in base-pairing and base-stacking (discussed in more detail below in §1.2.1).[72,75–79,94] T-jump studies of DNA conducted by Tokmakoff and co-workers has shown that short double-stranded DNA oligonucleotides undergo strand melting on the timescale of 10 – 30 μs, but with apparent "end-fraying" of the strand termini on the order of 70 – 100 ns,[78,79] indicating an unzipping of the strands from the ends toward the middle.

Studies by Stancik, et al investigating the dynamics of small RNA tetraloops using T-jump IR, indicated that during melting, W-C base-pairing is disrupted before base-stacking, and that for UNCG loops, when N=A, the loop structure around the adenine melts slower than either base-pairing or base-stacking.[25,80] This led

to the conclusion that the earliest phase of hairpin refolding involve base-stacking, with the reordering of the stacks relative to their neighbours until they are properly oriented for the chain to collapse into the base-paired stem. This is contrary to other studies which have determined that an initial correctly positioned base-pairing contact must be made, before stabilisation by the first base-stack facilitates the zipping of the stem.[95]

T-jump has also been a powerful tool for observing interactions between DNA and small binding molecules. Studies examining the change in DNA behaviour in the presence of a minor groove-binding ligand, Hoechst 33258, have shown that ligand binding results in an increase in the timescale of melting by up to a factor of 8 longer and suppresses the early pre-melting disruption of the structure, such as end-fraying.[75] It was also found to disrupt the stability of GC bases adjacent to the ligand binding sites.[72]

One of the challenges of T-jump measurements is that while the initial T-jump can be induced over nanoseconds and the response of a system monitored such as DNA strand melting, the solvent cooling typically takes on the order of 10s to 100s of milliseconds or longer.[77–80,96] This is too slow to capture native refolding dynamics seen on millisecond timescales.[92] As a consequence, the dynamics of refolding that would occur during a rapid cooling, or T-drop, is calculated from the rate of the melting extracted from fitting the T-jump kinetics and by using the equilibrium rate constant from the static temperature melting curves, typically using a two state model.[77–80] However, similar to melting, DNA folding is also not well described by a two state-transition, instead following "Anti-Arrhenius" behaviour where a negative activation energy is observed.[77] This is known to occur under certain conditions for reversible multi-step transitions, where each step independently follows Arrhenius behaviour.[97] For nucleic acid

folding this is due to the many potential semi-folded or incorrectly folded intermediate states on a "rugged" potential energy landscape.[81,95] As a consequence, refolding represents a significant loss of entropy, leading to a decreasing folding rate with increasing temperature.[97]

Considering the importance of the folding process for DNA and RNA to adopt the correct conformations necessary to perform their biological roles, characterising their folding behaviour would be a significant step toward understanding how these complex biomolecules carry out their biological roles. Direct measurement has been achieved with the use of a fast T-drop, using a thermally conductive sapphire substrate with light-absorptive nanocoating and a microfluidic device and fluorescent tetramethylrodamine labels.[98] While incredibly useful in some contexts, labels can influence the dynamics they are there to measure, and require additional preparation and measurements to interrogate multiple sites.[91] One of the aims of this work is to develop a label-free method for capturing melting and refolding dynamics of oligonucleotides in a single measurement.

1.2 Infrared Absorption Spectroscopy

In this thesis, the focus will be on the use of infrared spectroscopy in the study of RNA and DNA dynamics. The vibrational modes present in nucleic acids that make this analysis possible arise as a product of the stretching and bending of the bonds between atoms. For non-linear molecules that contain N atoms, there are 3N – 6 independent vibrational modes. These vibrational normal modes are independent, synchronous oscillations of atoms or groups of atoms that may be excited without exciting other normal modes or resulting in translation or rotation of the molecule. When the motion of the normal mode results in a

change in the dipole moment as a function of the vibrational coordinate, this satisfies the gross selection rule for IR activity.[99]

When a molecule is excited into a higher vibrational state by an infrared light source, it creates a coherent linear superposition of the vibrational ground and first excited states.[100] The time dependence of this superposition is called the molecular response, *R(t)*. The oscillating electric field of the light synchronises the excited vibrational modes, which when considered across many similarly excited molecules results in a non-equilibrium macroscopic polarisation, *P(t)*, of the charges in a sample. When the incident light does not carry a strong electric field strength, the macroscopic polarization can be approximated by the linear contribution of the macroscopic polarization function: [101,102]

$$\mathrm{P(t)} = \epsilon_0 \chi^{(1)} E(t) \qquad 1.1$$

The macroscopic polarization is determined by the permittivity of free space, ϵ_0, the first order electric susceptibility, $\chi^{(1)}$, and the oscillating electric field, *E(t)*, but is phase shifted by π/2 compared to the electric field.[101,102] This then decays over time as dictated by the molecular response, where the relationship between the linear response function, E' and macroscopic polarization is:[100]

$$\mathrm{P^{(1)}}(t) = \int_0^\infty \mathrm{d}t_1 E'(\mathrm{t}-\mathrm{t}_1) R^{(1)}(t) \qquad 1.2$$

This induced polarization gives rise to its own oscillating electromagnetic field, resulting in an emitted field signal (or free induction decay) from the sample in the same wavevector-matched direction as the polarization but with a further phase shift of π/2 [100,103] This leads the incident and emitted light to destructively interfere, leading to the observed decrease in light at that frequency, the absorption.

The macroscopic polarisation is measured with the intent to calculate the molecular response, which is characteristic of the molecule under examination. The polarization and the emitted field decay over what is called the dephasing time, T_2, and any population of the first excited state would relax back to the ground state with a vibrational lifetime of T_1, such that:

$$\frac{1}{T_2} = \frac{1}{2T_1} + \frac{1}{T_2^*} \qquad 1.3$$

Here T_2^* is pure dephasing time, a result of rapid fluctuations in the oscillation frequency of a molecule, resulting in additional loss of coherence and decay in the polarization. The dephasing rate is responsible for the natural, homogeneous linewidths in the infrared spectrum, with faster dephasing represented by broader linewidths. However, variation in the environment, for example through differences in H-bonding of solvent to solute molecule, leads to a range of vibrational frequencies for a specific mode in excited state molecules being populated, causing a shift in the spectrum and the overlap of these similar peaks to broaden the linewidth as well. This is called the inhomogenous linewidth.

Each vibrational mode has a characteristic frequency, though some may be degenerate, and if they are IR active this results in a characteristic absorption peak or band in the infrared spectrum. This frequency is determined by the reduced mass of the constituent atoms and the force constant, which is determined by the bond strengths. The intensity of the absorbance observed is described by the Beer-Lambert Law:

$$A = \varepsilon_\lambda \mathrm{cl} \qquad 1.4$$

The absorbance is directly proportional to the wavelength dependent molar extinction (ε_λ) coefficient, the concentration of the analyte in solution (c) and

the pathlength through the solution (l). The extinction coefficient of a given mode is dependent on the derivative of the dipole moment with respect to the vibrational coordinate.[104]

1.2.1 IR spectroscopy of Nucleic Acids

The vibrational modes of nucleic acids bases appear in the infrared spectral region from 1500 – 1800 cm^{-1} (Fig. 1.7 (a,b)), with distinct modes apparent for each type of nucleotide base. [58] This is the primary region of interest in studies conducted in this thesis as these modes are sensitive to the base-pairing and base-stacking transitions that occur during the melting process. For observation of the base-modes, deuterated water is used as the solvent, due to water having a large absorption peak at 1640 cm^{-1} corresponding to the HOH bending mode. The DOD bending mode appears at 1250 cm^{-1}.

In order to gain an understanding of the IR spectra of DNA it is useful to look at the features of the GC modes and the AT modes separately, as their base regions overlap to a significant degree and they report on distinct constituents of the molecule.

The spectra for the AT 10-mer shows characteristic peaks for double stranded (Fig. 1.7, (a) blue) and single stranded DNA (Fig. 1.7 (a), red). Spectra were assigned using notation consistent with studies by Cho et al and Banyay et al. [57-59] The transition dipole moment vectors for the non-paired base vibrational modes as determined by calculations conducted by Cho et al are shown below each spectrum. [59,60]

When single stranded, four peaks are apparent in the AT spectrum (Fig. 1.7, (a). The first is a relatively small peak seen at 1575 cm^{-1} which corresponds to a C=N

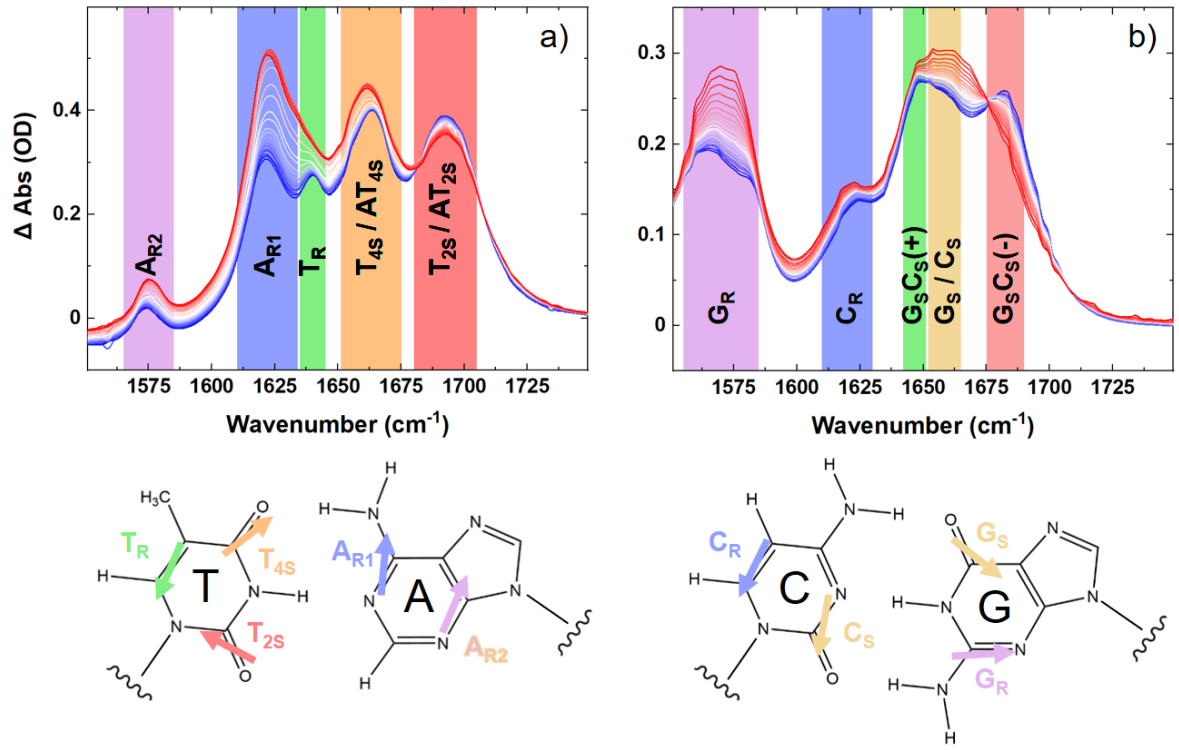

Figure 1.7: *DNA IR Absorption spectra showing vibrational modes present in double stranded DNA at 20°C (blue) through the melting transition to single stranded DNA at 80 °C (red) for a) AT modes for an AT DNA 10-mer and b) GC modes for a TNCG DNA hairpin with a four GC base paired stem. Respective localisation of non-paired base transition dipoles as calculated by Cho et al shown below each spectrum.*[60]

vibrational mode of the adenine ring, notated as A_{R2}.The most intense mode in the AT spectrum at 1622 cm^{-1} is a second distinct adenine ring mode, A_{R1}. At 1660 cm^{-1} the thymine C4=O stretching mode T_{4S} is observed, while the C2=O stretching mode T_{2S} appears at 1692 cm^{-1}.

On duplexation, there is a sharp decrease in the intensity of the A_{R1} mode observed. This change in intensity has been ascribed to suppression of the absorbance of the ring vibrational mode, due to the decrease in its extinction coefficient. This is caused by the change in the transition dipole moment as a consequence of base-stacking, with the formation of an extended π-system as

the π-orbitals of stacked base rings in a helix overlap.[58,105] This is similar to the hyperchromic effect observed in the UV-Vis spectrum and makes the A_{R1} useful for tracking the extent of duplexation with respect to AT pairs.

Upon forming the double stranded duplex, the hydrogen bonds between the pairs result in the individual base modes of the two bases coupling to produce distinct normal modes that are delocalised over the base-pair and provide specific markers of double stranded nucleic acids.[60] In general coupling occurs as a consequence of direct mechanical linking between localized normal modes by covalent bonds or electrostatic interaction through space. A consequence of this is that the oscillation of one influences the other.

In the AT spectrum, base pairing of the strands results in the appearance of the coupled adenine and thymine stretching modes, AT_{4S} and AT_{2S}, with AT_{4S} decreased in intensity compared to T_{4S}. The thymine ring vibrational mode, T_R does not change on strand association, but becomes visibly distinct from the A_{R1} mode which obscures it in single stranded conformation.[58]

In Fig. 1.7, (b), only three distinct GC modes are clearly visible in the single stranded form, the guanine C=N ring vibrational mode G_R at 1575 cm^{-1}, the cytosine ring mode, C_R at 1620 cm^{-1} and single peak at 1660 cm^{-1} where the peaks of the cytosine and guanine stretching modes (C_S and G_S) overlap. On base pairing, a characteristic decrease is seen in the G_R due to base stacking, similar to the A_{R2} mode for the AT base pair. Similarly, the G_R can be used as a proxy to determine the degree of duplexation in GC pairs. By contrast, no significant change is seen in the C_R with base-pairing or stacking. Hydrogen-bonding causes the C_S and G_S modes to become redshifted, but more significantly, the two modes strongly couple to form two new normal modes, the symmetric and

asymmetric stretching modes, $G_SC_S(+)$ and $G_SC_S(-)$. These are seen at 1646 cm^{-1} and 1686 cm^{-1} respectively.

While there are differences between the spectra of RNA and DNA with regard to frequency and relative intensity of different modes, they are very similar with equivalent mode assignments. As such, any relevant differences between RNA and DNA spectra will be highlighted in context in Chapters 3 – 5.

1.3 Temperature-Jump IR Spectroscopy

The primary focus of work presented in this thesis is the conformational dynamics of RNA and DNA hairpins. In order to probe these nanosecond to millisecond nucleic acid dynamics it is necessary to introduce a perturbation, such as a T-jump, to create a sustained non-equilibrium state over the timescale of interest and initiate dynamic conformational response that can be observed in real time. T-jumps have been induced through the use of capacitive discharge through a solution resulting in microsecond temperature rise times,[106] but more typically an IR laser pulse is tuned to a solvent band, resulting in a nanosecond temperature rise time.[107] This system is then monitored by a series of IR probe pulses to measure the nucleic acid's IR absorption spectrum at discrete time points to follow its dynamic response.

For the work conducted in this thesis, the T-jump was initiated by tuning the T-jump pump to the OD stretching mode at 2750 cm^{-1} in deuterated water. This causes an initial excitation in the OD stretch. Due to the extended hydrogen bonding network that evolves on a sub-picosecond timescale and delocalization of the vibrational modes of water, the system relaxes rapidly through a series of energy transfer pathways.

In general, coupling of the vibrational modes in water allows the energy to flow from the excited state into a combination of lower frequency modes, where

conservation of energy dictates the sum of the relaxation frequencies matches the excitation frequency.[108,109] Experiments in water have shown that the hydrogen bonding interactions cause the frequency of the OH stretch to red-shift whereas the HOH bending frequency blue-shifts. A consequence of this is that the OH oscillator shifts towards resonance with the bending overtone transition, which results in a Fermi resonance between the v=1 state for hydrogen-bonded OH oscillators and the v=2 state of the bending mode.[110] This is the preferred relaxation route of the OH stretch. However, the hydrogen bonds to neighbouring water molecules also result in the stretching mode being delocalized over multiple molecules resulting in intermolecular energy transfer between stretching modes as well.[111] Similarly, calculations of the HOH bending vibration have shown a rapid of transfer energy via 2:1 Fermi resonance with the hindered rotation (libration) of the excited water molecule, which accounts for the majority of its energy flow. The rest is transferred to the four immediate neighbouring waters via hydrogen bond facilitated coupling of the bending modes.[112] By these intra and intermolecular pathways, which have been observed in D_2O as well,[109,112–114] the T-jump pump energy is rapidly dissipated to lower frequency modes, with randomized orientation and dispersed throughout the water network on a sub-picosecond timescale toward an entropically favoured distribution.[115] With the excess energy originating from the initial excitation spread over the whole ensemble of water molecules, a macroscopic temperature is established.[113]

The T-jump is induced by a nanosecond pulse, much longer than the relaxation of water or any coherent phenomena and therefore can be considered as a simple heater. Using this heating effect however enables a powerful method for investigating the dynamics processes of nucleic acid system. The aims of this thesis are to deepen our understanding of a specialised RNA UNCG tetraloop

motif using T-jump IR spectroscopy to study its melting and refolding dynamics. Chapter 2 will outline the methodology and experimental apparatus employed throughout the work presented in this thesis. In Chapter 3 a method will be developed for capturing the real time, native refolding dynamics of nucleic acids via a combined T-jump/Drop spectroscopic methods and will be used to probe the comparative melting behaviour of analogous RNA UNCG and DNA TNCG tetraloops. Chapter 4 will extend this work to explore the sequences dependence of the RNA and DNA hairpins stems on their melting and folding dynamics. Chapter 5 will employ T-jump with 2D-IR probing to investigate the melting pathways of the RNA and DNA hairpins. The data supporting the research presented in this thesis is openly available from the Research Data York repository.[116]

1.4 References

1 K. Sneppen and G. Zocchi, *Physics in Molecular Biology*, Cambridge University Press, Cambridge, 2005.

2 Li. Allison, *Fundamental molecular biology*, Wiley-Blackwell, 2021, vol. 35.

3 J. Hizver, H. Rozenberg, F. Frolow, D. Rabinovich and Z. Shakked, *Proc. Natl. Acad. Sci. U. S. A.*, 2001, **98**, 8490–8495.

4 L. R. Comolli, N. B. Ulyanov, A. M. Soto, L. A. Marky, T. L. James and W. H. Gmeiner, *Nucleic Acids Res.*, 2002, **30**, 4371–4379.

5 T. Zhou, Y. Tsybovsky, J. Gorman, M. Rapp, G. Cerutti, G. Y. Chuang, P. S. Katsamba, J. M. Sampson, A. Schön, J. Bimela, J. C. Boyington, A. Nazzari, A. S. Olia, W. Shi, M. Sastry, T. Stephens, J. Stuckey, I. T. Teng, P. Wang, S.

Wang, B. Zhang, R. A. Friesner, D. D. Ho, J. R. Mascola, L. Shapiro and P. D. Kwong, *Cell Host Microbe*, 2020, **28**, 867–879.

6 J. G. Yodh, M. Schlierf and T. Ha, *Q. Rev. Biophys.*, 2010, **43**, 185–217.

7 G. Houlihan, S. Arangundy-Franklin and P. Holliger, *Acc. Chem. Res.*, 2017, **50**, 1079–1087.

8 P. A. Jeggo, L. H. Pearl and A. M. Carr, *Nat. Rev. Cancer*, 2016, **16**, 35–42.

9 I. Tinoco, *Annu. Rev. Phys. Chem.*, 2002, **53**, 1–15.

10 T. Schlake, A. Thess, M. Fotin-Mleczek and K. J. Kallen, *RNA Biol.*, 2012, **9**, 1319–1330.

11 N. Pardi, M. J. Hogan, F. W. Porter and D. Weissman, *Nat. Rev. Drug Discov.*, 2018, **17**, 261–279.

12 P. A. Morcos, *Biochem. Biophys. Res. Commun.*, 2007, **358**, 521–527.

13 D. Kang, Y. Lee and J. S. Lee, *Cancers (Basel).*, 2020, **12**, 2699.

14 F. Crick and J. Watson, *Nature*, 1953, **171**, 737–738.

15 P. Yakovchuk, E. Protozanova and M. D. Frank-Kamenetskii, *Nucleic Acids Res.*, 2006, **34**, 564–574.

16 M. Zacharias, *J. Phys. Chem. B*, 2020, **124**, 10345–10352.

17 F. Kilchherr, C. Wachauf, B. Pelz, M. Rief, M. Zacharias and H. Dietz, *Science*, 2016, **353**, 1116.

18 I. Anosova, E. A. Kowal, M. R. Dunn, J. C. Chaput, W. D. V. Horn and M. Egli, *Nucleic Acids Res.*, 2016, **44**, 1007–1021.

19 H. Schindelin, M. Zhang, R. Bald, J. P. Fürste, V. A. Erdmann and U. Heinemann, *J. Mol. Biol.*, 1995, **249**, 595–603.

20 J. E. Johnson and C. G. Hoogstraten, *J. Am. Chem. Soc.*, 2008, **130**, 16757–16769.

21 R. T. Batey, R. P. Rambo and J. A. Doudna, *Angew. Chemie - Int. Ed.*, 1999, **38**, 2326–2343.

22 C. R. Woese, S. Winker and R. R. Gutell, *Proc. Natl. Acad. Sci. U. S. A.*, 1990, **87**, 8467–8471.

23 V. P. Antao, S. Y. Lai and I. Tinoco, *Nucleic Acids Res.*, 1991, **19**, 5901–5905.

24 J. M. Blose, D. J. Proctor, N. Veeraraghavan, V. K. Misra and P. C. Bevilacqua, *J. Am. Chem. Soc.*, 2009, **131**, 8474–8484.

25 A. L. Stancik and E. B. Brauns, *J. Phys. Chem. B*, 2013, **117**, 13556–13560.

26 J. H. A. Nagel, C. Flamm, I. L. Hofacker, K. Franke, M. H. De Smit, P. Schuster and C. W. A. Pleij, *Nucleic Acids Res.*, 2006, **34**, 3568–3576.

27 V. Baumruk, C. Gouyette, T. Huynh-Dinh, J. S. Sun and M. Ghomi, *Nucleic Acids Res.*, 2001, **29**, 4089–4096.

28 L. D'Ascenzo, Q. Vicens and P. Auffinger, *Nucleic Acids Res.*, 2018, **46**, 7989–7997.

29 D. W. Staple and S. E. Butcher, *PLoS Biol.*, 2005, **3**, 0956–0959.

30 F. M. Jucker, H. A. Heus, P. F. Yip, E. H. M. Moors and A. Pardi, *J. Mol. Biol.*, 1996, **264**, 968–980.

31 H. A. Heus and A. Pardi, *Science*, 1991, **253**, 191–194.

32 G. Varani, *Annu. Rev. Biophys. Biomol. Struct*, 1995, 379–404.

33 C. Cheong, G. Varani and I. Tinoco, *Nature*, 1990, **346**, 680–682.

34 L. Dascenzo, F. Leonarski, Q. Vicens and P. Auffinger, *RNA*, 2017, **23**, 259–269.

35 G. Varani and W. H. McClain, *EMBO Rep.*, 2000, **1**, 18–23.

36 M. R. Giese, K. Betschart, T. Dale, C. K. Riley, C. Rowan, K. J. Sprouse and M. J. Serra, *Biochemistry*, 1998, **37**, 1094–1100.

37 V. P. Antao and I. Tinoco, *Nucleic Acids Res.*, 1992, **20**, 819–824.

38 G. Varani, C. Cheong and I. Tinoco, *Biochemistry*, 1991, **30**, 3280–3289.

39 B. Fürtig, C. Richter, J. Wöhnert and H. Schwalbe, *ChemBioChem*, 2003, **4**, 936–962.

40 K. Wuthrich, *Acta Crystallogr. - Sect. D Biol. Crystallogr.*, 1995, **51**, 249–270.

41 E. P. Nikonowicz and A. Pardi, *Nature*, 1992, **355**, 184–186.

42 D. Marion, *Mol. Cell. Proteomics*, 2013, **12**, 3006–3025.

43 W. P. Aue, E. Bartholdi and R. R. Ernst, *J. Chem. Phys.*, 1976, **64**, 2229–2246.

44 L. Baronti, I. Guzzetti, P. Ebrahimi, S. Friebe Sandoz, E. Steiner, J. Schlagnitweit, B. Fromm, L. Silva, C. Fontana, A. A. Chen and K. Petzold, *Nature*, 2020, **583**, 139–144.

45 K. C. Engman, P. Sandin, S. Osborne, T. Brown, M. Billeter, P. Lincoln, B. Nordén, B. Albinsson and L. M. Wilhelmsson, *Nucleic Acids Res.*, 2004, **32**, 5087–5095.

46 C. R. Cantor and I. Tinoco, *J. Mol. Biol.*, 1965, **13**, 65–77.

47 M. M. Warshaw and I. Tinoco, *J. Mol. Biol.*, 1966, **20**, 29–38.

48 E. Emanuele, D. Markovitsi, P. Millié and K. Zakrzewska, *ChemPhysChem*, 2005, **6**, 1387–1392.

49 M. D'Abramo, C. L. Castellazzi, M. Orozco and A. Amadei, *J. Phys. Chem. B*, 2013, **117**, 8697–8704.

50 J. SantaLucia and D. Hicks, *Annu. Rev. Biophys. Biomol. Struct.*, 2004, **33**, 415–440.

51 H. T. Allawi and J. SantaLucia, *Nucleic Acids Res.*, 1998, **26**, 2694–2701.

52 M. Menger, F. Eckstein and D. Porschke, *Biochemistry*, 2000, **39**, 4500–4507.

53 Y. You, A. V. Tataurov and R. Owczarzy, *Biopolymers*, 2011, **95**, 472–486.

54 G. R. Bishop and J. B. Chaires, *Curr. Protoc. Nucleic Acid Chem.*, 2002, **11**, 7.11.1-7.11.8.

55 C. R. Cantor, M. M. Warshaw and H. Shapiro, *Biopolymers*, 1970, **9**, 1059–1077.

56 V. I. Ivanov, L. E. Minchenkova, E. E. Minyat, M. D. Frank-Kamenetskii and A. K. Schyolkina, *J. Mol. Biol.*, 1974, **87**, 817–833.

57 J. M. Benevides and G. J. Thomas, *Nucleic Acids Res.*, 1983, **11**, 5747–5761.

58 M. Banyay, M. Sarkar and A. Gräslund, *Biophys. Chem.*, 2003, **104**, 477–488.

59 C. Lee, K. H. Park and M. Cho, *J. Chem. Phys.*, 2006, **125**, 114508.

60 C. Lee and M. Cho, *J. Chem. Phys.*, 2006, **125**, 114509.

61 G. Hithell, D. J. Shaw, P. M. Donaldson, G. M. Greetham, M. Towrie, G. A.

Burley, A. W. Parker and N. T. Hunt, *J. Phys. Chem. B*, 2016, **120**, 4009–4018.

62 F. B. Howard, J. Frazier and H. . T. Miles, *Proc. Natl. Acad. Sci. U. S. A.*, 1969, **64**, 451–458.

63 E. Taillandier and J. Liquier, *Methods Enzymol.*, 1992, **211**, 307–335.

64 M. Falk, K. A. Hartman and R. C. Lord, *J. Am. Chem. Soc.*, 1963, **85**, 391–394.

65 H. T. Miles and J. Frazier, *Biochem. Biophys. Res. Commun.*, 1963, **14**, 21–28.

66 E. M. Bruening, J. Schauss, T. Siebert, B. P. Fingerhut and T. Elsaesser, *J. Phys. Chem. Lett.*, 2018, **9**, 583–587.

67 A. T. Krummel and M. T. Zanni, *J. Phys. Chem. B*, 2006, **110**, 13991–14000.

68 G. Hithell, G. A. Burley, P. M. Donaldson, N. T. Hunt, A. W. Parker, M. Towrie, K. Wynne, G. M. Greetham and M. González-Jiménez, *Phys. Chem. Chem. Phys.*, 2017, **19**, 10333–10342.

69 L. A. I. Ramakers, G. Hithell, J. J. May, G. M. Greetham, P. M. Donaldson, M. Towrie, A. W. Parker, G. A. Burley and N. T. Hunt, *J. Phys. Chem. B*, 2017, **121**, 1295–1303.

70 A. T. Krummel, P. Mukherjee and M. T. Zanni, *J. Phys. Chem. B*, 2003, **107**, 9165–9169.

71 R. Fritzsch, P. M. Donaldson, G. M. Greetham, M. Towrie, A. W. Parker, M. J. Baker and N. T. Hunt, *Anal. Chem.*, 2018, **90**, 2732–2740.

72 X. X. Zhang, S. L. Brantley, S. A. Corcelli and A. Tokmakoff, *Commun. Biol.*,

2020, **3**, 1–9.

73 B. P. Fingerhut and T. Elsaesser, in *Springer Series in Optical Sciences*, Springer Verlag, 2019, vol. 226, pp. 171–195.

74 J. Schauss, F. Dahms, B. P. Fingerhut and T. Elsaesser, *J. Phys. Chem. Lett.*, 2019, **10**, 238–243.

75 R. Fritzsch, G. M. Greetham, I. P. Clark, L. Minnes, M. Towrie, A. W. Parker and N. T. Hunt, *J. Phys. Chem. B*, 2019, **123**, 6188–6199.

76 J. Dale, C. P. Howe, H. Toncrova, R. Fritzsch, G. M. Greetham, I. P. Clark, M. Towrie, A. W. Parker, T. C. McLeish and N. T. Hunt, *Phys. Chem. Chem. Phys.*, 2021, **23**, 15352–15363.

77 R. J. Menssen and A. Tokmakoff, *J. Phys. Chem. B*, 2019, **123**, 756–767.

78 P. J. Sanstead, P. Stevenson, A. Tokmakoff and A. Tokmako, *J. Am. Chem. Soc.*, 2016, **138**, 11792–11801.

79 P. J. Sanstead and A. Tokmakoff, *J. Phys. Chem. B*, 2018, **122**, 3088–3100.

80 A. L. Stancik and E. B. Brauns, *Biochemistry*, 2008, **47**, 10834–10840.

81 H. Ma, D. J. Proctor, E. Kierzek, R. Kierzek, P. C. Bevilacqua and M. Gruebele, *J. Am. Chem. Soc.*, 2006, **128**, 1523–1530.

82 R. K. Nayak, O. B. Peersen, K. B. Hall and A. Van Orden, *J. Am. Chem. Soc.*, 2012, **134**, 2453–2456.

83 R. Narayanan, L. Zhu, Y. Velmurugu, J. Roca, S. V. Kuznetsov, G. Prehna, L. J. Lapidus and A. Ansari, *J. Am. Chem. Soc.*, 2012, **134**, 18952–18963.

84 M. Novakovic, G. L. Olsen, G. Pintér, D. Hymon, B. Fürtig, H. Schwalbe and L. Frydman, *Proc. Natl. Acad. Sci. U. S. A.*, 2020, **117**, 2449–2455.

85 E. A. Dethoff, K. Petzold, J. Chugh, A. Casiano-Negroni and H. M. Al-Hashimi, *Nat. 2012 4917426*, 2012, **491**, 724–728.

86 A. L. Lieblein, J. Buck, K. Schlepckow, B. Fürtig and H. Schwalbe, *Angew. Chemie - Int. Ed.*, 2012, **51**, 250–253.

87 H. Shi, A. Rangadurai, H. Abou Assi, R. Roy, D. A. Case, D. Herschlag, J. D. Yesselman and H. M. Al-Hashimi, *Nat. Commun.*, 2020, **11**, 1–14.

88 M. Hairong, C. Wan, A. Wu and A. H. Zewail, *Proc. Natl. Acad. Sci. U. S. A.*, 2007, **104**, 712–716.

89 K. Sarkar, K. Meister, A. Sethi and M. Gruebele, *Biophys. J.*, 2009, **97**, 1418–1427.

90 J. L. Fiore and D. J. Nesbitt, *Q. Rev. Biophys.*, 2013, **46**, 223–264.

91 K. Sarkar, D. A. Nguyen and M. Gruebele, *RNA*, 2010, **16**, 2427–2434.

92 D. J. Proctor, H. Ma, E. Kierzek, R. Kierzek, M. Gruebele and P. C. Bevilacqua, *Biochemistry*, 2004, **43**, 14004–14014.

93 S. V. Kuznetsov, C. C. Ren, S. A. Woodson and A. Ansari, *Nucleic Acids Res.*, 2008, **36**, 1098–1112.

94 C. P. Howe, G. M. Greetham, B. Procacci, A. W. Parker and N. T. Hunt, *J. Phys. Chem. Lett.*, 2022, **13**, 9171–9176.

95 W. Zhang and S. J. Chen, *Biophys. J.*, 2006, **90**, 765–777.

96 B. Ashwood, N. H. C. Lewis, P. J. Sanstead and A. Tokmakoff, *J. Phys. Chem. B*, 2020, **124**, 8665–8677.

97 L. E. Revell and B. E. Williamson, *J. Chem. Educ.*, 2013, **90**, 1024–1027.

98 M. E. Polinkovsky, Y. Gambin, P. R. Banerjee, M. J. Erickstad, A. Groisman

and A. A. Deniz, *Nat. Commun.*, 2014, **5**, 5737.

99 P. Atkins and J. de. Paula, *Physical Chemistry*, Oxford University Press, 2009.

100 P. Hamm and M. T. Zanni, *Concepts and Methods of 2D Infrared Spectroscopy*, Cambridge University Press, Cambridge, 2011.

101 P. W. Atkins and R. S. Friedman, *Molecular Quantum Mechanics*, Oxford University Press, Oxford, 1996, vol. 9.

102 S. Suresh, A. Ramanand, D. Jayaraman and P. Mani, *Rev. Adv. Mater. Sci.*, 2012, **30**, 175–183.

103 M. D. Fayer, Ed., *Ultrafast Infrared Vibrational Spectroscopy*, CRC Press, Boca Raton, 1st edn., 2013.

104 B. S. Galabov and T. Dudev, Eds., in *Vibrational Spectra and Structure*, Elsevier Science, 1996, vol. 22, pp. 77–140.

105 G. W. Doorley, M. Wojdyla, G. W. Watson, M. Towrie, A. W. Parker, J. M. Kelly and S. J. Quinn, *J. Phys. Chem. Lett.*, 2013, **4**, 2739–2744.

106 J. Kubelka, *Photochem. Photobiol. Sci.*, 2009, **8**, 499–512.

107 N. T. Hunt, *Chem. Soc. Rev.*, 2009, **38**, 1837–1848.

108 V. M. Kenkre, A. Tokmakoff and M. D. Fayer, *J. Chem. Phys.*, 1994, **101**, 10618–10629.

109 T. Steinel, J. B. Asbury, J. Zheng and M. D. Fayer, *J. Phys. Chem. A*, 2004, **108**, 10957–10964.

110 K. Ramasesha, L. De Marco, A. Mandal and A. Tokmakoff, *Nat. Chem. 2013 511*, 2013, **5**, 935–940.

111 W. B. Carpenter, J. A. Fournier, R. Biswas, G. A. Voth and A. Tokmakoff, *J. Chem. Phys.*, 2017, **147**, 84503.

112 R. Rey, F. Ingrosso, T. Elsaesser and J. T. Hynes, *J. Phys. Chem. A*, 2009, **113**, 8949–8962.

113 R. Rey and J. T. Hynes, *Phys. Chem. Chem. Phys.*, 2012, **14**, 6332–6342.

114 F. Ingrosso, R. Rey, T. Elsaesser and J. T. Hynes, *J. Phys. Chem. A*, 2009, **113**, 6657–6665.

115 W. J. Smit and H. J. Bakker, *J. Chem. Phys.*, 2013, **139**, 204504.

116 C. P. Howe and N. T. Hunt, *Research Data York*, DOI:10.15124/a02eb3ff-5245-463f-b434-8e16e6e5b653.

2. Methods

This chapter outlines the methods and instrumentation used for all experimental work in this thesis: the preparation of samples, linear FT-IR absorption spectroscopy and T-Jump IR Spectroscopy. The T-jump experiments used the STFC Central Laser facility's ULTRA laser system, located in the Research Complex at Harwell at the Rutherford Appleton Laboratories. Additional methods for specific experiments are discussed in context in Chapter 3 - 5.

2.1 Sample Preparation

Salt-free, lyophilised DNA and RNA oligomer sequences were purchased from Eurogentec. Deuterium oxide, potassium phosphate (mono and dibasic salts) and deuterium chloride were purchased from Sigma-Aldrich and used without modification. The alternating stem sequence of the tetraloops was chosen to minimise base pair slippage.

All DNA and RNA samples were prepared to a strand concentration of 10 mM, below the threshold for duplex formation, in 1 M deuterated phosphate buffer (pD 6.8).[1] For all spectroscopy measurements a 15-20 µl aliquot was placed in a temperature controlled cell (Harrick, ±1 ^{o}C) equipped with CaF_2 windows. A 50 µm PTFE spacer defined the path length for IR absorption measurements while a 12 µm path length was used for all other measurements.

2.2 Infrared Absorption Spectroscopy

All RNA and DNA samples throughout Chapter 3 – 5 were characterised in the solution phase by Infrared absorption spectroscopy. Infrared absorption spectra were collected using a commercial Bruker Vertex 70 Fourier Transform (FT)-IR spectrometer. Using a black-body emitter, such as a silicon carbide filament, a "Globar", a broadband incoherent mid-infrared light is generated across a frequency range from 600-7000 cm^{-1} (Fig 2.1). This light is collimated and passed through a Michelson Interferometer.[2]

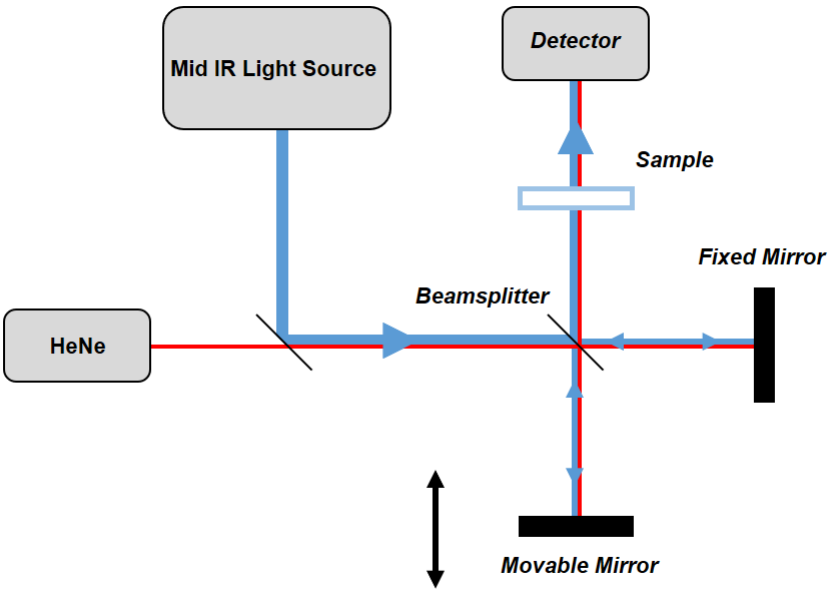

Figure 2.1: *Schematic diagram of an IR spectrometer using a Michelson interferometer and a Helium-Neon (HeNe) reference laser.*

The interferometer consists of a beam splitter which divides the beam and directs it down two paths, one with a fixed length and the other varied by a mirror moving along the path of the beam to adjust the pathlength. The beams are then recombined at the beamsplitter, with the relative phase between them controlled by their different roundtrip times, then passed through the sample and directed onto a mercury cadmium telluride (MCT) point detector. At

different positions of the mirrors, different frequencies of light will interfere constructively and destructively, leading to a time dependent frequency profile.

To Fourier transform this signal from the time domain to the frequency domain to generate the frequency axis, the pathlength at any given time must be known. To determine this, a Helium-Neon (HeNe) laser is used, which generates a beam with a constant wavelength and is passed through the interferometer, creating an interference pattern that is maximised when the variable pathlength is at multiples of the wavelength, allowing the pathlength to be precisely determined. A separate measurement without a sample is used as a reference to determine the intensity of the absorbance (A) of the sample by the change in the intensity of light before and after the sample, the ratio of the probe and reference signals, I and I_0 respectively:[2]

$$A = -\log\frac{I}{I_0} \qquad 2.1$$

The intensity of transmitted light varies as described by the Beer-Lambert Law (see Chapter 1, eq. 1.4). Concentrations and pathlengths used in the work described in Chapters 3 - 5 were controlled so that the oligonucleotide absorbance were maintained at a value of 1 or less, with 1 representing absorption of 90% of the light. Spectra were collected with a resolution of 1 cm^{-1}, with 20 scans. The spectra were solvent corrected.

2.3 Temperature-Jump Infrared Spectroscopy

In Chapters 3 and 4 work is presented using time resolved spectroscopy with a laser induced temperature-jump (T-jump) to study the different melting dynamics of RNA UNCG and DNA TNCG tetraloop hairpins. Chapter 5 extends

this work, using T-jump with 2D-IR probing to observe the changes during hairpin melting (discussed further in Chapter 5)

The Temperature-jump measurements were conducted with the STFC Central Laser Facility's ULTRA B spectrometer, employing the time resolved multiple probe spectroscopy (TRMPS) method.[3–7] In order to conduct this experiment, a T-jump is initiated by IR laser pulse tuned to a D_2O vibrational mode. This vibrational excitation relaxes on a sub-picosecond timescale,[3] dispersing its energy into other solvent modes in the system. This causes a perturbation of the target molecule such as nucleic acids dissolved in the solution, equivalent to a macroscopic increase in the temperature. The response of the nucleic acids to the perturbation can then be captured in real-time using a series of broadband infrared probes over the millisecond cooling time of the T-jump.[3,4]

2.3.1 T-jump Pump

The generation of the T-jump pump pulses is based on a neodymium-doped yttrium aluminium garnet (Nd:YAG) pumped optical parametric oscillator (OPO). The Nd:YAG laser (Fig. 2.2) is a Bright Solutions Wedge XB, short- pulse Q-switched diode-pumped solid-state laser and generates pulses centred at 1064 nm with a pulse duration of <1.5 ns, a pulse energy of 4 mJ and a 1 kHz pulse repetition rate. 3.2 mJ of this is focused onto a Covesion MOPO-1.0–2.0 magnesium oxide doped, periodically poled lithium niobate (MgO:PPLN) crystal in a ~4 cm singly resonant OPO cavity.[3] By adjusting the temperature of the crystal, the pump pulses generated were tuned to the high frequency edge of the symmetric O-D stretch of the D_2O solvent at 3600 nm (2750 cm^{-1}) with a pulse energy of 70 μJ and pulse duration of 4 ns.

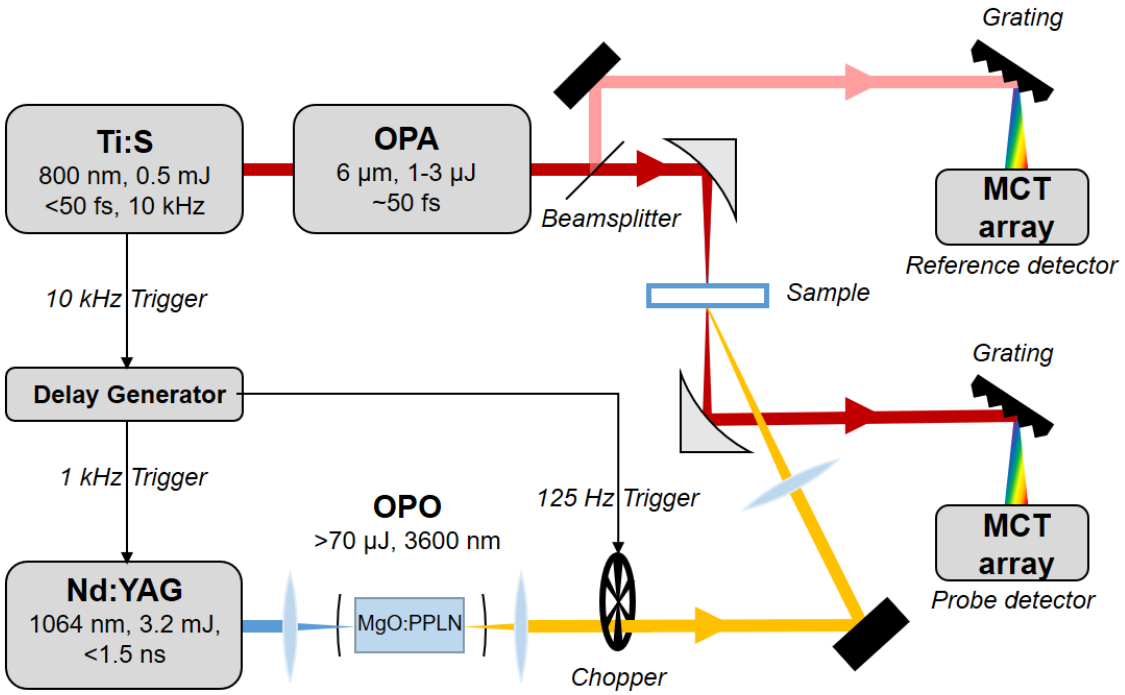

Figure 2.2: *Schematic of the ULTRA T-jump instrumental set up.*[3] *The T-jump pulses are followed by a series of probes at 100 µs intervals to scan the processes in the perturbed system over the millisecond cooling time. A delay generator is used to enable probing with nanosecond time resolution.*

2.3.2 T-jump Mid-IR Probe

To generate the probe pulses for the T-jump experiment, a Ti:S oscillator is employed to produce pulses to seed the amplifier. The pulses generated are centred at 800 nm with a bandwidth of 50 nm with a pulse width of 20 fs at a repetition rate of 65 MHz. The amplifier (Coherent Legend Elite Duo) is a titanium doped sapphire (Ti:S) chirped-pulse amplifier that increases the energy of the pulses generated by the oscillator. The seed pulses are first ”chirped” by using a grating to temporally disperse the pulse by frequency in order to temporarily lower the peak intensity of the pulses to avoid damage to the optics. This results in an increase in pulse duration from 20 fs to 1 ps.

The stretched pulse is then regeneratively amplified using a cryogenically cooled Ti:S crystal, pumped by a Q-switched neodymium-doped yttrium lithium fluoride (Nd:YLF) laser at 1053 nm. Electro-optic modulators (Pockels cells) at either end of the amplification cavity control ingress and egress of the cavity. Here a single pulse resonates within the cavity, passing through the Ti:S gain medium until maximum amplification is achieved. The amplified pulse is then ejected from the cavity, where a grating reverses the chirp, producing a 5W output pulse with a 10 kHz repetition rate and a pulse duration of <50 fs.

The amplified pulse are directed into an optical parametric amplifier (OPA) with difference frequency generation (DFG). This converts the 800 nm light from the amplifier into the mid-IR range. The 800 nm pulse from the amplifier is split into a high intensity pump pulse (P_1) and a low intensity pulse (P_2). P_2 is then further split, into high (P_3) and low intensity (P_4) pulses. P_4 is then incident on a sapphire plate, generating a white light continuum. P_3 is used to pump a β-barium borate (BBO) crystal, with the white light pulses temporally broadened before being used to seed the optical parametric amplification process in the crystal, producing a signal and idler pulses with frequencies such that:

$$\omega_{\text{pump}} = \omega_{\text{signal}} + \omega_{\text{idler}} \qquad 2.2$$

In order to select the desired output signal frequency, the delay between the white light and pump pulses and the relative angle of the BBO plate are adjusted. This signal pulse is then directed to a second BBO crystal, which is pumped by P_1 collinearly to the signal pulse. This amplifies the signal according to the relationship:

$$\omega_{\text{pump}} + \omega_{\text{signal}} = 2\omega_{\text{signal}} + \omega_{\text{idler}} \qquad 2.3$$

This also produces an idler pulse collinear to the signal. Through difference frequency mixing of the signal and idler in a silver thiogallate ($AgGaS_2$) crystal, light at ~6 μm with a bandwidth of >300 cm^{-1} full width at half maximum (FWHM) is produced with a 1-3 μJ pulse energy.

2.3.3 T-jump Spectrometer

With the pump and probe pulses both generated, the timing between them needs to be controlled. The pump pulse was synchronised with the probe using a Stanford DG645 to produce an adjustable phase computer-controlled delay. The 10 kHz probe trigger timing was divided down to 1 kHz, triggering the Nd:YAG laser, with a further divide down to 125 Hz to trigger the pump-beam chopper.

The probe light is focused on a sample with a 50 μm spotsize, overlapped with the larger 300 μm pumped sample area. The transmitted probe light is dispersed onto 128-channel mercury cadmium telluride (MCT) array detectors. Before the sample, 10% of the probe light is split and dispersed onto a 32 - channel MCT detector as a reference to correct for any shot to shot variation in the probe, improving the stability of absorption measurements.

With a pump repetition rate of 125 Hz after optical chopping, and a probe repetition rate of 10 kHz, for every T-jump (Fig. 2.3, yellow pulse) a series of 80 probes are measured at intervals of 100 μs (Fig. 2.3, red pulses). The total measurement window between consecutive T-jumps is 8 ms. To measure data at any chosen time points between T-jump pump pulse, a series of electronic delays are introduced between the pump and the probes, so the first probe is recorded at x nanoseconds, with the nth probe (n between 2-80) recorded at x + (100 μs)*n. Schematically, in Fig. 2.3 this would appear as a shift of the red

probes pulses compared to the yellow T-jump pulse. After averaging, the next electronic delay is measured, shifting the probes to a new temporal position. A result of this probing method is that on a log time axis often used to visualise the dynamics, data points are seen to cluster at intervals of 100 µs.

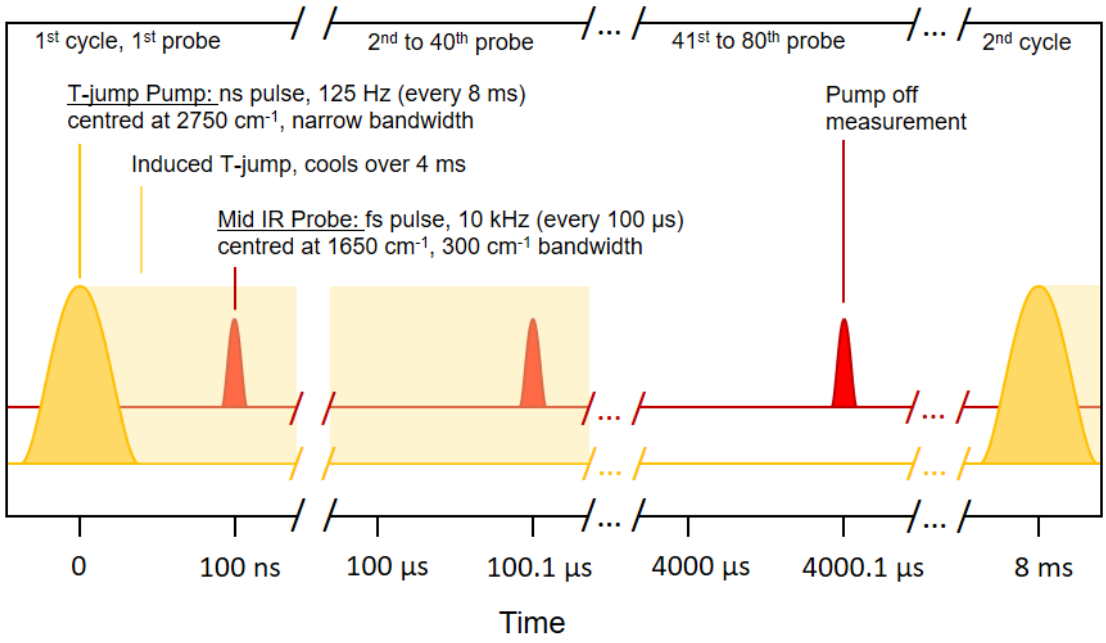

Figure 2.3: *Schematic diagram of the pulse timings used for TRMPS T-jump pump IR probe spectrometer. Yellow indicates ns T-jump pump pulse (repetition rate 125 Hz) while red indicates the mid-IR probe pulses (repetition rate 10 kHz).*

The T-jump is seen to cool and return to the baseline within 4 ms, which allows the final 4 ms to be used as a pump off measurement. This is particularly useful because electromagnetic frequency (EMF) noise is produced by the high voltage Pockels cells switch of the pump laser that causes a time dependent ringing in the data, which appears as high frequency noise at regular time intervals.[3] The

subtraction of the pump-off allows this ringing to be corrected as the same EMF is generating in the pump-on and pump-off measurements.

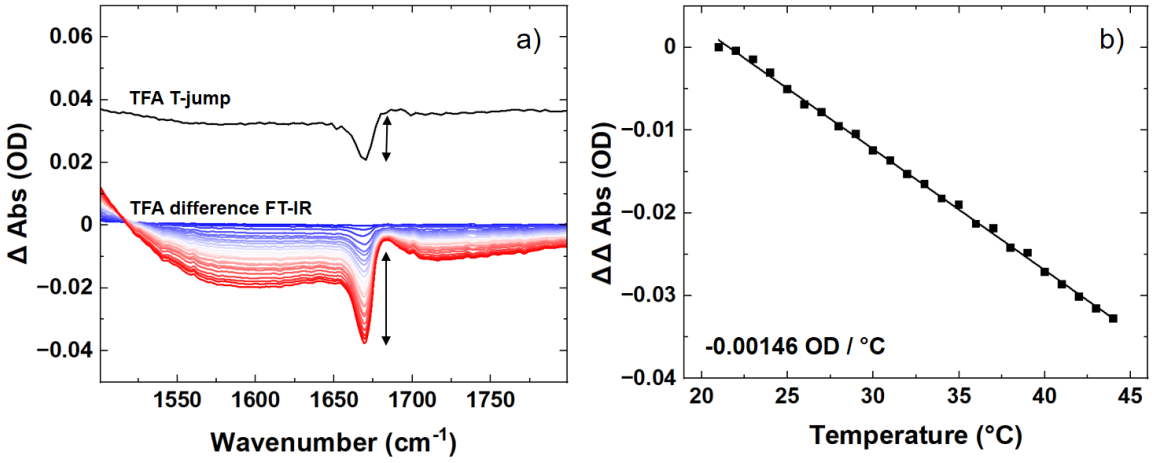

Figure 2.4: *Calibration of the T-jump using TFA, a) comparison of T-jump response for T_0=20 °C at 200 ns to the TFA FT-IR difference spectra from 20 – 44 °C, b) temperature dependence of the TFA T-jump peak – peak response for the 1670 cm^{-1} asymmetric carboxyl stretching mode.*

To measure the magnitude of the temperature rise induced by the T-jump pulse used to capture dynamics within a desired temperature window, the T-jump needs to be calibrated by comparison to a reference with a known response. For this purpose Trifluoroacetic acid (TFA) was used as a calibrant due to its small structure and direct hydrogen bonding interaction with the solvent leading to fast dynamics that then follow the slower cooling dynamics of the solvent. It also has an intense, narrow linewidth IR signal at 1670 cm^{-1} (Fig. 2.4 (a)), the asymmetric carboxylate stretch of TFA.[3,4] With increasing temperature, this signal decreases in intensity linearly with a shift to higher frequency. Thus the peak to peak difference can be plotted as a function of temperature as shown in Figure 2.4 (b). Using this the magnitude of the T-jump temperature rise can

be determined by measuring the maximum TFA signal in the T-jump and comparing it against the temperature dependence of the TFA peak. For the experiments described in this thesis, the T-jump was calibrated to provide a 10 °C increase in temperature.

2.4 References

1 A. L. Stancik and E. B. Brauns, *J. Phys. Chem. B*, 2013, **117**, 13556–13560.

2 P. Atkins and J. de. Paula, *Physical Chemistry*, Oxford University Press, 2009.

3 G. M. Greetham, I. P. Clark, B. Young, R. Fritsch, L. Minnes, N. T. Hunt and M. Towrie, *Appl. Spectrosc.*, 2020, **74**, 720–727.

4 R. Fritzsch, G. M. Greetham, I. P. Clark, L. Minnes, M. Towrie, A. W. Parker and N. T. Hunt, *J. Phys. Chem. B*, 2019, **123**, 6188–6199.

5 J. Dale, C. P. Howe, H. Toncrova, R. Fritzsch, G. M. Greetham, I. P. Clark, M. Towrie, A. W. Parker, T. C. McLeish and N. T. Hunt, *Phys. Chem. Chem. Phys.*, 2021, **23**, 15352–15363.

6 C. P. Howe, G. M. Greetham, B. Procacci, A. W. Parker and N. T. Hunt, *J. Phys. Chem. Lett.*, 2022, **13**, 9171–9176.

7 L. Minnes, G. M. Greetham, D. J. Shaw, I. P. Clark, R. Fritzsch, M. Towrie, A. W. Parker, A. J. Henry, R. J. Taylor and N. T. Hunt, *J. Phys. Chem. B*, 2019, **123**, 8733–8739.

3. Measuring RNA UNCG Tetraloop Refolding Dynamics Using Temperature-Jump/Drop Infrared Spectroscopy

3.1 Introduction

Here we use T-jump IR spectroscopy to measure the melting and refolding dynamics of a 12 nucleotide RNA sequence comprising a UACG tetraloop and four base pair double-stranded GC stem, shown in figure 3.1. This was compared to an equivalent DNA (TACG) sequence. Due to the all GC stem, these sequences are notated as RNA_{all} and DNA_{all}.

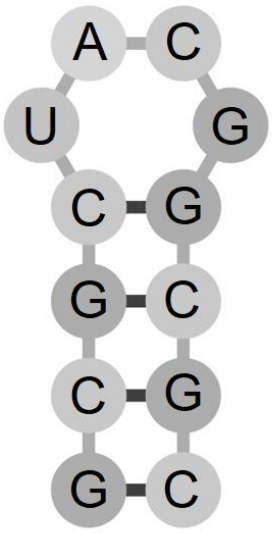

Figure 3.1: *Schematic representation of the base sequence of the RNA UACG tetraloop RNA_{all}. The 5' end is located at the bottom left of the structure.*

The melting characteristics of UNCG loops have been determined previously by static spectroscopy experiments,[1,2] but the folding dynamics have yet to be resolved. To address this important issue, this chapter outlines a new method to provide a sample cooling rate that is fast enough to allow direct observation of strand refolding. Using IR spectroscopy as a probe of vibrational modes of the

nucleic acid bases provides real-time insight into stem base-pairing and stacking. Although DNA hairpins are rare biologically, comparing RNA tetraloop dynamics with those of an analogous DNA sequence reveals inherent differences between molecules with overtly similar molecular structures.

3.2 Experimental

The RNA and DNA oligomer sequences 5'-GCGC(XACG)GCGC-3' (RNA: X = U; DNA: X = T; "(...)" indicates the loop position) were prepared as outlined in §2.1.

FT-IR Absorption measurements of the sequences were conducted over the temperature range from 20 – 80 °C through the melting transition with points taken at intervals of 3 °C.

The Temperature-jump used the STFC Central Laser Facility's ULTRA spectrometer T-jump method.[3–6] A detailed description of the T-jump set up is described in Chapter 2, §2.3. Measurements were taken at a range of starting temperatures, T_0, over the temperature range from 20-80 °C to determine the temperature dependence of hairpin dynamics.

3.3 Results and Analysis

3.3.1 FT-IR Absorption Spectroscopy

The IR absorption spectra of RNA_{all} and DNA_{all} hairpins (Fig. 3.2) were collected to characterise and assign their different base vibrational modes and observe their behaviour over the transition from fully double stranded, to partially melted and finally fully melted spectra.

Key spectroscopic signatures are observed in the region from 1550 to 1700 cm^{-1}. In the case of RNA_{all} (Fig. 3.2 (a)) and DNA_{all} (Fig. 3.2 (b)), the spectra contain contributions from predominantly guanine and cytosine modes, but the loop contains an adenine and a uracil in RNA_{all} (thymine in DNA_{all}) and due to the unpaired bases present in the loop structure, even in a fully folded hairpin configuration, both modes assigned to W-C paired bases and unpaired bases are present.[7–9]

Comparing the spectra of RNA_{all} and DNA_{all}, the RNA_{all} $G_SC_S(+)$ and $G_SC_S(-)$ modes (Fig. 3.2 (a,b), green and red panels) are more intense in the low temperature spectra (blue traces). When the sequences are melted (red traces), RNA_{all} also shows higher intensity G_S and C_S modes, but broadly the DNA and RNA high temperature spectra show the same features, although the lineshape at 1660 cm^{-1} in DNA_{all} is broader. Some of the $G_SC_S(-)$ is still visible at high temperature in the RNA_{all} spectrum as it is not completely melted at 80 °C, due to its higher melting point.

For both hairpins, the A_{R2} mode is obscured by the more prominent G_R (Fig. 3.2, purple panel), so to avoid confusion, the A_{R1} mode, visible at 1622 cm^{-1} (Fig. 3.2, blue panel), will be referred to simply as A_R moving forward.

Melting curves for the RNA_{all} and DNA_{all} hairpins were obtained by plotting the absorbance of the G_R mode as a function of temperature (Fig. 3.3) and fitting (lines) to a sigmoidal function:

$$A(T) = A_2 + \frac{A_1 - A_2}{1 + e^{(T - T_m)/d_T}} \qquad 3.1$$

Where A_1 and A_2 are the starting and maximum absorbance, T_m, the melting point at 50% of the rise in absorbance, and the temperature constant d_T.

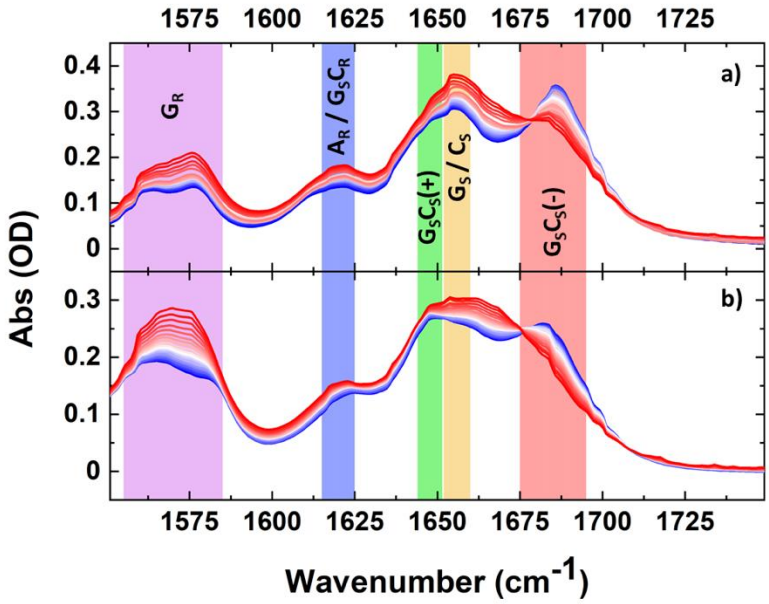

Figure 3.2: *FT-IR spectra for a) RNA_{all} and b) DNA_{all} from 20 to 80 °C (blue-red). Prominent vibrational modes are indicated by coloured panels: G_R (1575 cm^{-1}, purple), A_R and G_SC_R (1620 cm^{-1}, blue), G_SC_S(+) (1648 cm^{-1}, green), G_S and C_S (1656 cm^{-1}, yellow) and G_SC_S(-) (1686 cm^{-1}, red).*[7–9] *Spectra have been solvent corrected.*

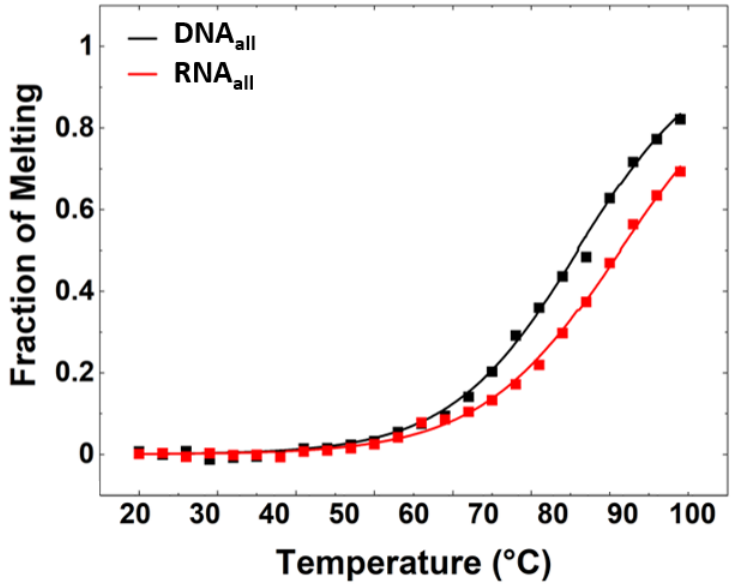

Figure 3.3: *The temperature dependent behaviour shown by FT-IR melting curves for RNA_{all} (red) and DNA_{all} (black) extracted from the G_R mode (1575 cm^{-1}) along with sigmoidal fits (lines). Data is normalised to the sigmoid fit maximum, A_2 (see Eq. 3.1), to show fraction of melting.*

For the fitting, the melting curves have been baseline corrected by subtracting the 1750 cm^{-1} behaviour, which is in a region with no spectral features, and subtracting a linear temperature baseline extrapolated from the 20 – 40 °C range where no melting occurs. A summary of the fitting parameters with standard errors is shown in Table 3.1.

Table 3.1: Summary of the Melting Curve Fitting

	Parameter	Value	Dependency
DNA_{all}	**A_1**	-2.18 ± 1.99 (10^{-4})	0.56
	A_2	0.039 ± 0.002	0.96
	T_m	76.0 ± 0.8	0.95
	d_T	8.2 ± 0.4	0.84
	R^2	0.998	
RNA_{all}	**A_1**	-0.86 ± 1.82 (10^{-4})	0.60
	A_2	0.048 ± 0.003	0.98
	T_m	81.3 ± 1.1	0.98
	d_T	8.8 ± 0.5	0.89
	R^2	0.998	

The binding energies of the hairpin stems are quite high, especially in the case of RNA_{all}, so the sequences did not achieve 100% melting. To account for this, the data was normalised to the maximum of the sigmoid fit, A_2, rather than the highest temperature data point to approximate the fraction of melting (Fig. 3.3). The melting curves clearly show the different temperature behaviour between the DNA_{all} and RNA_{all} hairpins, with the inflection point of the curves indicating the melting point of DNA determined to be 76 ± 1 °C, while for RNA it is 81 ± 1 °C.

The thermodynamic parameters for the melting transition were obtaining from the melting curves using the Van't Hoff equation:

$$\frac{d}{dT}\ln K_{eq} = \frac{\Delta H}{RT^2} \qquad 3.2$$

The Van't Hoff equation relates the enthalpy of melting, ΔH, at a given temperature T, with $\frac{d}{dT}$, the slope of the melting curve at temperature T, the ideal gas constant R and the equilibrium rate constant, K_{eq}. Since these hairpins are monomolecular and below the threshold concentration for duplex formation, the melting behaviour of the sequences was concentration independent as established by thermodynamic studies on similar UNCG hairpins.[15] By combining the Van't Hoff equation with the equation for Gibbs free energy, the entropy, ΔS, is obtained by:

$$\ln K_{eq} = -\frac{\Delta H}{RT} + \frac{\Delta S}{R} \qquad 3.3$$

The calculated thermodynamic parameters of melting are shown in Table 3.2. Uncertainty is indicated by standard errors.

Table 3.2: Thermodynamic parameters obtained via Van't Hoff analysis

		RNA_{all}	DNA_{all}	
Van't Hoff	**ΔH**	118.3 ± 9.4	124.2 ± 8.2	$kJmol^{-1}$
	ΔS	334 ± 27	356 ± 23	$JK^{-1}mol^{-1}$
	ΔG*	14.8 ± 12.6	13.9 ± 10.8	$kJmol^{-1}$

**ΔG calculated at 37°*C

3.3.2 Temperature-Jump IR Spectroscopy

It is important for any time resolve methods that the means of initiation is faster than the process under investigation. For the temperature rise seen in the T-jump this has already been established, with a temperature rise time of nanoseconds compared to oligonucleotide melting dynamics seen over 100s of nanoseconds to microseconds.[4,5] In order to develop a T-jump method with

sufficiently fast cooling of the solvent to enable the native dynamics of RNA and DNA to be captured, an understanding of the rates of cooling of the T-jump and the refolding rates of the oligonucleotides needs to be established. A study of T-jump dynamics determined cooling of a 10 °C T-jump with a pathlength of 10 μm occurs over 1 ms, with the recovery time to half maximum shown to be dependent on the square of the pathlength.[5] As a result reducing the pathlength significantly increases the cooling rate. This was determined to be driven by the greater thermal conductivity of the windows compared to the water. With a greater pathlength, the surface area in contact with the windows is lower compared to the volume heated by the T-jump. The change in heat distribution through the water is also less significant compared to loss from the windows due to the pump beam having a diameter six times the probe beam diameter.

Calculations on RNA folding show that, while the folding rate is variable depending on sequence, the lifetime of refolding was limited to 1 ms.[10,11] Studies of small double stranded RNA hairpin ribozyme structures in a rapid mixing experiment observed complete refolding had occurred in the time range between 1 and 10 ms. [12,13] Using different methods and oligonucleotides, these methods are not directly comparable, however they suggest that with the right parameters, the cooling rate could be increased to be faster than the refolding rate of the hairpins under investigation.

Solvent Cooling Dynamics

To determine the lifetime of D_2O cooling, T-jump measurements were conducted using TFA to observe the cooling rate of D_2O after the T-jump (Fig. 3.4). This was done with pathlengths of 6 and 12 μm in order to observe the effect of pathlength on the cooling rate. The temporal profile of a vibrational

mode of the TFA at 1670 cm^{-1} was fitted using a bi-exponential (fitting discussed in greater detail below) to extract the lifetime of cooling. At 6 µm, the lifetime of cooling was found to be 39 ± 2 µs, while at 12 µm it was 149 ± 2 µs. This is consistent with the squared relationship between pathlength and cooling rate as determined by simulation studies.[5]

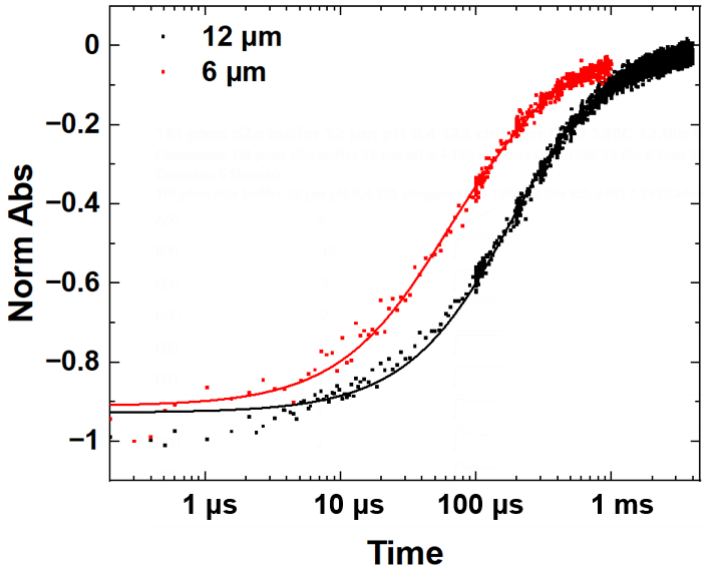

Figure 3.4: *Normalised T-jump kinetics (dots) for TFA with pathlengths of 6 (red) and 12 µm (black). Kinetics were fit (lines) using a double exponential to determine the lifetime of cooling.*

To verify whether this was sufficiently fast cooling to capture the refolding of RNA_{all} and DNA_{all}, T-jump kinetics measurements of RNA_{all} at 75 °C and DNA_{all} at 70 °C were taken with pathlengths of 6 and 12 µm (Fig. 3.5). These temperatures represent the maximum change observed over the melting transition for each. The kinetics for the G_R mode were extracted and fit to a triple exponential to determine the lifetime of cooling/refolding.

The time constants for the decay of the G_R signal for RNA_{all} were determined to be 210 ± 4 and 283 ± 3 µs for 6 and 12 µm respectively, while for DNA_{all} they were found to be 122 ± 1 and 190 ± 3 µs. These times are both longer than the

observed cooling rate of the solvent. While a larger separation between the solvent cooling and nucleic acid dynamics was observed at 6 µm, a path length of 12 µm was chosen for further measurements due to the larger signal sizes and lower error in the pathlength, improving data quality and providing greater consistency between samples. This was important to allow more reliable comparison between RNA and DNA.

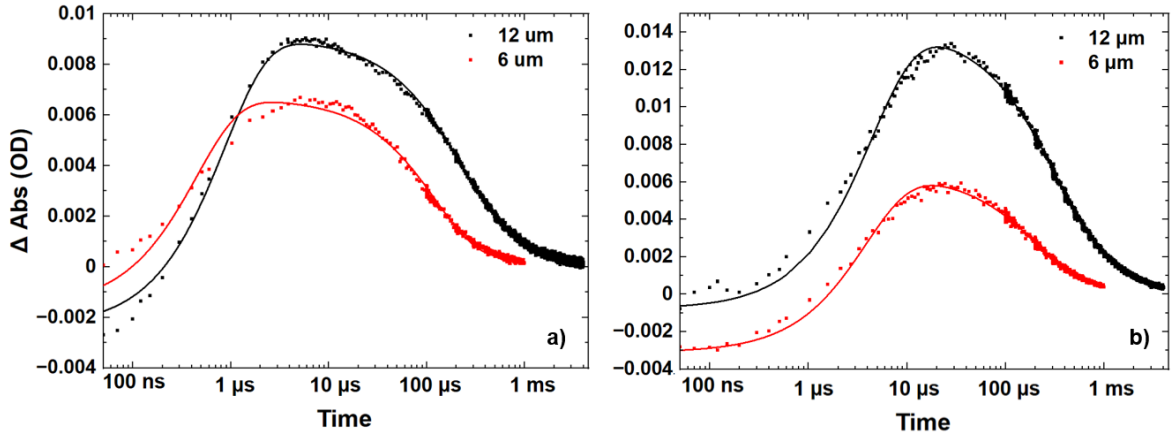

Figure 3.5: *T-jump kinetics (dots) of the G_R mode for a) DNA_{all} at 70 °C and b) RNA_{all} at 75 °C with pathlengths of 6 (red) and 12 µm (black). Kinetics were fit (lines) using a triple exponential to determine the lifetime of cooling.*

Temperature Dependent Dynamics of RNA_{all} and DNA_{all}

Having determined that the natural heating and cooling processes of the solvent in response to the T-jump with a 12 µm pathlength is faster than the rise and and decay rates of the DNA_{all} and RNA_{all} hairpins signals, T-jump measurements of RNA_{all} and DNA_{all} were conducted over a range of temperatures through the melting transition. This was done to establish conclusively that both melting and refolding conformational dynamics of RNA_{all} and DNA_{all} were being captured and to characterise their temperature dependent behaviour. The results of the T-jump spectroscopy measurements are shown in Fig. 3.6 (c, d). The T-jump

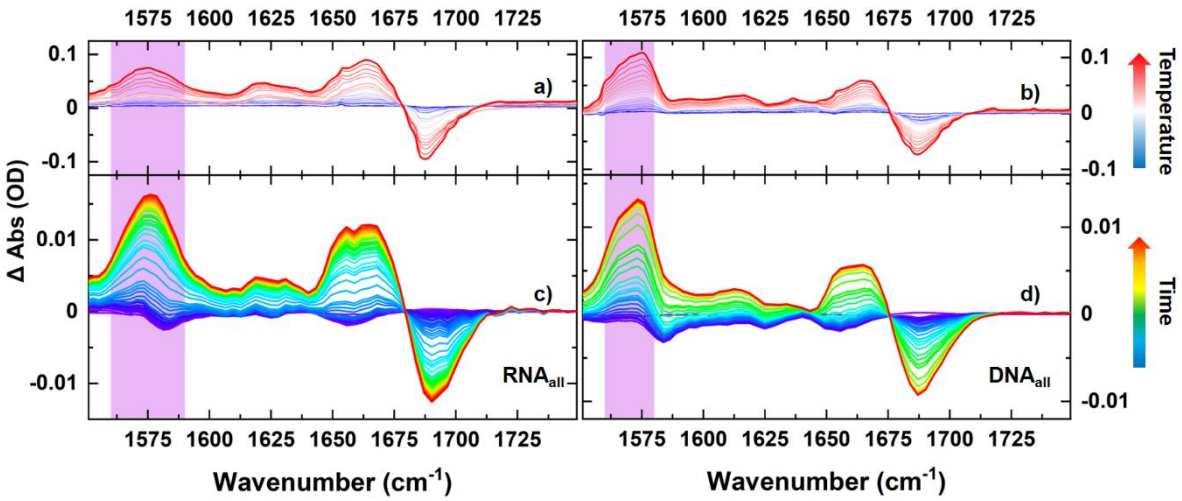

Figure 3.6: *Comparison of FT-IR and T-jump IR measurements, showing solvent corrected FT-IR difference spectra as a function of T relative to the spectrum at 20 °C of a) the RNA_{all} and b) DNA_{all} tetraloops. The colour scale runs from 20-80 °C (blue-red) with the G_R mode highlighted. Panels c) and d) show the T-jump IR spectra of RNA_{all} and DNA_{all} respectively (T_0=T_m-5 °C). Time delays from 1 ns (blue) to peak signals (red) at 20 µs (RNA_{all}) and 6 µs (DNA_{all}) are displayed. The T-jump data are represented as T-jump on – T-jump off difference spectra; the increase in amplitude of the G_R band upon melting of the double stranded GCGC stem appears as a positive peak highlighted by the purple panel. The negative spectral features at very early time in the T-jump are due to a fast hydrogen bonding rearrangement. For visual clarity only, T-jump spectra have been baseline corrected.*

induces a 10 °C temperature rise, so the data shown were obtained with an initial T_0 five degrees below the melting temperature of the sample ($T_0 = T_m – 5$ ˚C) to observe the greatest change over the T-jump.

The spectra are represented as T-jump on minus T-jump off difference spectra and show changes in band positions and intensities that occur over a period of up to 20 microseconds following the T-jump (blue-red). For both RNA_{all} and

DNA_{all} samples, the spectral changes following the T-jump were observed to be identical to those in the difference IR absorption spectra (Fig. 3.6 (a, b)) and therefore are assigned to melting of the stem of the loop as the hairpin responds to the change in sample temperature.

Comparing the maximum intensity change of the G_R band in the T-jump experiment to the change in the same band obtained via FT-IR spectroscopy over the same temperature range, showed that the extent of melting achieved for RNA_{all} and DNA_{all} hairpins was 83% and 98% respectively.[2] This establishes that the hairpin samples undergo virtually all of the melting expected over the temperature range during the T-jump experiment.

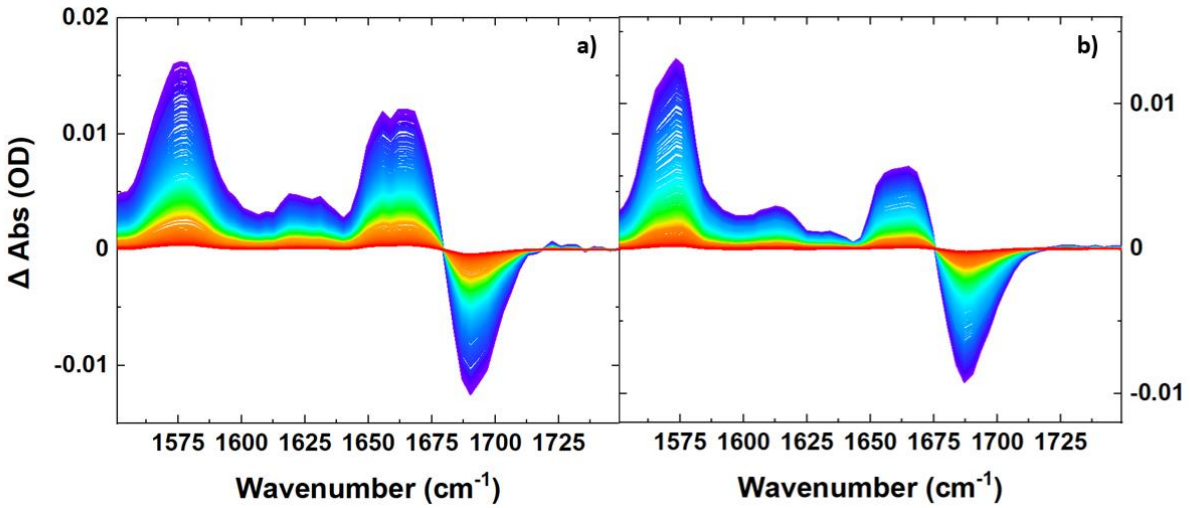

Figure 3.7*: T-jump spectra for a) the RNA_{all} hairpin at a T_0 of 75 °C showing changes due to refolding, showing data obtained at T-jump-probe delay times between 20 µs and 4 ms (blue to red) and b) DNA_{all} refolding from 6 µs and 4 ms (blue to red). It is clear that the refolding process reverses the spectral changes that occur during melting. For visual clarity, T-jump spectra have been baseline corrected by subtracting the 1750 cm^{-1} time-dependent response.*

On longer timescales the signals decayed to the baseline as the sample cooled following the T-jump, with the spectral changes for both RNA_{all} (Fig. 3.7, (a)) and DNA_{all} (Fig. 3.7, (b)) showing the reverse trend compared to the melting response, indicating refolding of the hairpins over the 4 ms T-jump cooling time.

To examine the dynamics of the hairpins following the T-jump, the time and temperature dependence of the G_R band was analysed (Fig. 3.8). In each case, the G_R band intensity was observed to rise, peaking near 20 µs and 6 µs for RNA_{all} (Fig. 3.8, b) and DNA_{all} (Fig. 3.8, e) hairpins respectively, before decaying on millisecond timescales. The relative temperature dependent solvent response was seen to be small compared to the response of the G_R (Fig. 3.8 (c))

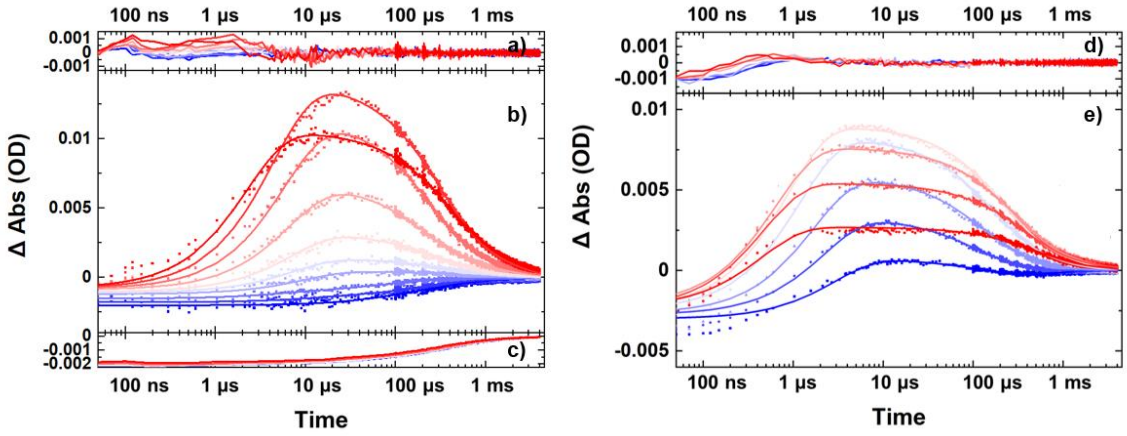

Figure 3.8. *The temperature- and time-dependence of the G_R band intensity at 1575 cm^{-1} for b) the RNA_{all} hairpin and e) the DNA_{all} hairpin. Data (b and e, dots) are shown from a T_0 of 50 °C to 85 °C (blue-red) at intervals of 5 °C through the hairpin melting transition, along with the triple exponential fits (b and e, lines) and residuals (a and d). c) The temperature dependent D_2O solvent contribution, shown to scale. No solvent correction was applied to RNA_{all} and DNA_{all} as the solvent contribution is small (~15% of the signal) and thus does not contribute materially to the temperature dependent behaviour.*

To further confirm that the process observed in the T-jump was following the same behaviour observed in the equilibrium FT-IR absorption measurements, the temperature dependence of the G_R maximum in the T-jump at 20 µs and 6 µs for RNA_{all} and DNA_{all} respectively was examined (Fig. 3.9). For comparison to the equilibrium temperature measurements, the T-jump temperature plotted is T_0 + 5 °C, the average temperature over the T-jump. The temperature corresponding to the maximum observed signal in the T-jump was 75 °C for DNA_{all} and 80 °C for RNA_{all}. In both instances this is within 1 °C of the T_m extracted from the equilibrium FT-IR melting curves establishing that T-jump and FT-IR data are probing the same physical process.

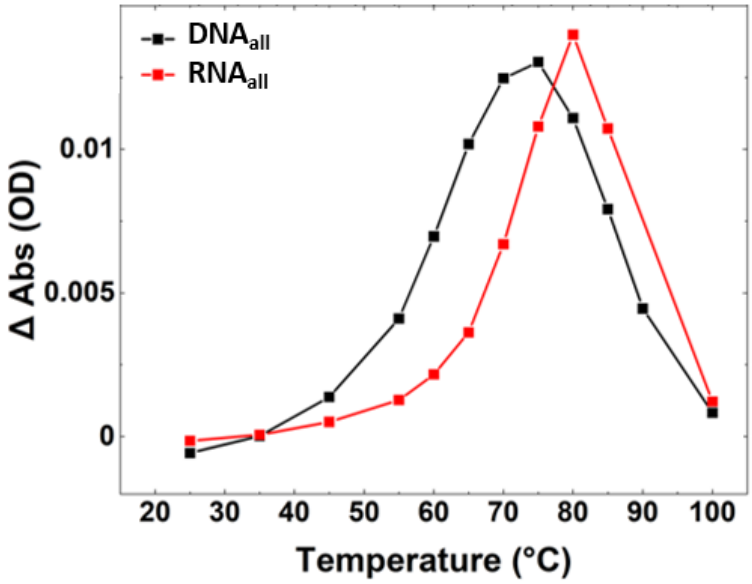

Figure 3.9: *The temperature dependence of the change in absorbance of the T-jump signal (T_0 + 5 °C) for RNA_{all} (red) and DNA_{all} (black) extracted from the G_R mode (1575 cm^{-1}) at the T-jump delay corresponding to the signal maximum (20 µs for RNA_{all}, 6 µs for DNA_{all}).*

Analysis of T-jump Dynamics

In order to extract the dynamic parameters, the lifetimes (τ) of melting and refolding, appropriate fitting methods must be applied that are reflective of the processes occurring over from T-jump. To determine the best approach, both double stretched exponential (Eq. 3.4) and triple exponential (Eq. 3.5) functions were tested by fitting the temporal dependence of the G_R band (Fig 3.10).

$$\Delta A(T) = A_0 + A_1 e^{-(\frac{T}{\tau_1})^{\beta_1}} + A_2 e^{-(\frac{T}{2})^{\beta_2}} \quad 3.4$$

$$\Delta A(T) = A_0 + A_1 e^{-(\frac{T}{\tau_1})} + A_2 e^{-(\frac{T}{\tau_2})} + A_3 e^{-(\frac{T}{\tau_3})} \quad 3.5$$

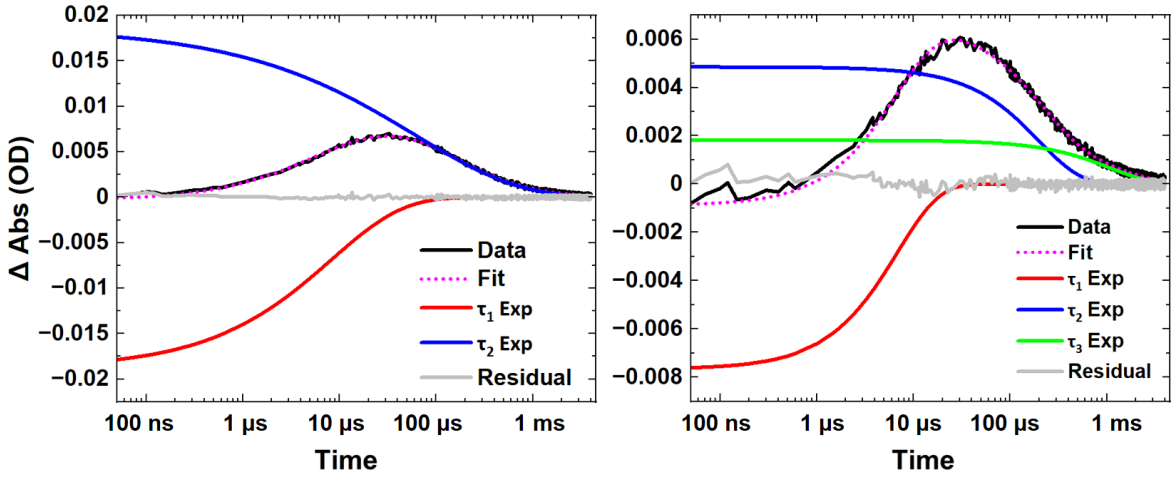

Figure 3.10: *Fitting of the G_R T-jump kinetics for RNA_{all} at 65 °C using, a) double stretched exponential and b) triple exponential. Data is shown in black, magenta dotted lines show the fit and red, blue and green traces are the individual exponential components of the fit. The residuals of the fit are shown in grey.*

The exponential stretching factor, β, is typically used to account for variations in the state of the ensemble of molecules being probed during the melting process.[4,14,15] Specifically in this instance, the greatest variation will be a consequences of the temperature gradient induced across the cell by the T-jump.[5] However, in these measurements the cooling is also relatively fast, with the temperature gradient reducing as the sample cools, which may result in the

stretching factor not being a good description. The fitting displays low values of β (Table 3.3), which results in stretching the exponentials along the time axis (Fig. 3.10 (a), red and blue traces) and thus the melting and refolding are not reliably resolved using this method.

In addition, although the double stretched exponential closely fit the data it often, though not consistently, converged on solutions with pre-exponential terms (A_1, A_2 in Eq. 3.4) unrelated to the signal sizes seen in the spectroscopy.

Table 3.3: Summary of the fitting of the G_R T-jump Kinetics of RNA_{all}

	Parameter	Value	Dependency
Double Stretched Exponential	A_0	1.34 ± 0.01 (10^{-4})	0.947
	A_1	-0.019 ± 0.001	0.999
	A_2	0.019 ± 0.001	0.999
	τ_1	8166 ± 115	0.945
	τ_2	60500 ± 7800	0.999
	β_1	0.58 ± 0.01	0.979
	β_2	0.41 ± 0.01	0.999
	R^2	0.995	
Triple Exponential	A_0	1.07 ± 0.08 (10^{-4})	0.950
	A_1	-0.008 ± 0.001	0.704
	A_2	0.005 ± 0.001	0.948
	A_3	0.002 ± 0.001	0.988
	τ_1	6800 ± 100	0.571
	τ_2	199300 ± 3000	0.955
	τ_3	1084300 ± 31600	0.984
	R^2	0.990	

By comparison the triple exponential consistently converged on fits with pre-exponential terms that were well matched to the signal sizes and lifetimes that were temporally separated (Fig. 3.10 (b)). The majority of the temporal profile was accounted for by two processes with lifetimes quantifying the rise (τ_1) with lifetimes on the order of 500 ns to 10 μs, and decay (τ_2) of the signal with lifetimes of 100-400 μs. The third exponential term (τ_3) was generally of low amplitude (< 25%) but was essential to achieve a good quality fit. It was found to be on the order of 800 μs, significantly longer than other dynamics. Based on

its magnitude, which was greater than the solvent signal (Fig. 3.8 (c)), and its non-constant temperature dependence (see Table 3.4 below), this contribution could not be attributed simply to solvent response or a cooling property of the cell. As such it may indicate a slower dynamic process that follows the initial refolding, but the small nature of the contribution precludes clear conclusions. The triple exponential fitting displayed improved dependency of the parameters and crucially the fitting was consistent across samples and temperatures and thus better suited to direct comparison between RNA and DNA. With better resolution of the melting and refolding and better reliability, the triple exponential was used to fit all nucleic acid dynamics.

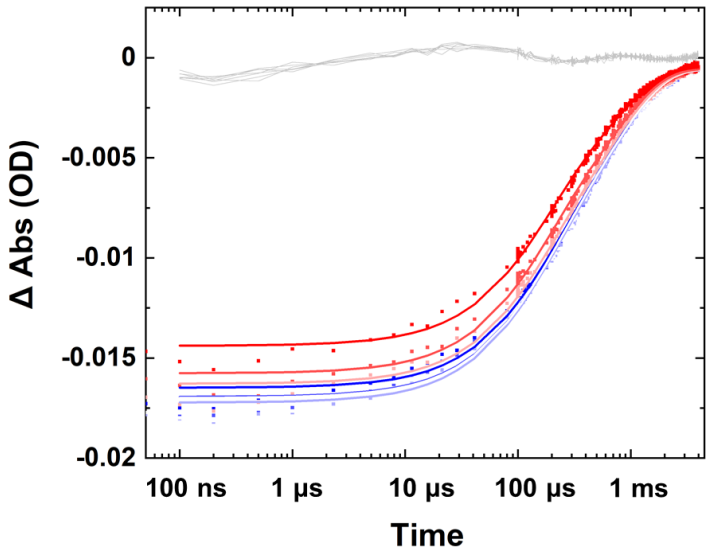

Figure 3.11: *Temperature dependent T-jump spectroscopy data showing the temporal response of the carboxyl stretching vibrational mode of trifluoroacetic acid (TFA, dots) along with fits to bi-exponential functions (lines), at T_0 values from 20 to 80 °C (blue to red). Fit residuals are shown as grey lines. Fits were conducted from 100 ns to avoid the initial T-jump rise; no baseline or solvent correction was performed.*

The temperature dependent cooling dynamics of the D_2O solvent were also measured (Fig. 3.11), for comparison with the RNA_{all} and DNA_{all} data. A bi-exponential function was used to mirror the approach for nucleic acid dynamics while accounting for the fact that the rise time of the T-jump in the solvent is essentially instantaneous, following the profile of the nanosecond laser pulse. For the TFA data, the second timescale recovered from the exponential function was equivalent to the small τ_3 contribution observed for RNA_{all} and DNA_{all} and so the major TFA relaxation timescale is used as an indicator of sample cooling time.

The results of the TFA fitting showed that the cooling timescale of the solvent was invariant with T_0. To confirm that these results were consistent in samples containing hairpins, the cooling dynamics of a band due to D_2O (1470 cm^{-1}) present in both the TFA calibration sample and in the hairpin samples were compared and found to display similar dynamics.

The fitting data for one set of T-jump experiments for RNA_{all}, DNA_{all} and the solvent dynamics (as determined by TFA) are shown in Table 3.4. Errors for the fitting parameters are <5%. These parameters are taken from the fits through the melting transition from a T_0 of 65-80 °C and 60-80 °C for RNA_{all} and DNA_{all} respectively, fitted over the entire time range from -20 ns to 4 ms. In order to focus on experiments where a significant proportion of hairpin melting occurred, data is shown over the temperature range where the maximum change in intensity of the G_R band was greater than 20% of the largest signal observed.

For further analysis, the timescales were averaged over three independent experimental measurements in order to determine the reproducibility and error between T-jump measurements.

Table 3.4: Fitting Parameters for RNA_{all}, DNA_{all} and TFA

	T (°C)	A_1 (OD)	A_2 (OD)	A_3 (OD)	τ_1 (ns)	τ_2 (µs)	τ_3 (µs)	A_0 (OD)
RNA_{all}	65	-0.0077	0.0049	0.0018	6860	199	1080	0.0142
	70	-0.0123	0.0083	0.0030	6120	220	1070	0.0236
	75	-0.0149	0.0102	0.0037	4520	283	1260	0.0287
	80	-0.0114	0.0070	0.0034	2260	350	1530	0.0216
DNA_{all}	60	-0.0099	0.0050	0.0017	1530	135	583	0.0167
	65	-0.0117	0.0062	0.0025	1120	169	716	0.0203
	70	-0.0118	0.0062	0.0026	765	220	920	0.0205
	75	-0.0097	0.0051	0.0017	570	298	1350	0.0164
	80	-0.0067	0.0034	0.0012	475	449	8370	0.0120
TFA	70		0.0067	0.0085		149	757	0.0005

It is noted that D_2O gives rise to a small, broadband solvent-dependent contribution to the data. In the analyses conducted this was not corrected for, but comparisons showed that applying this correction led to changes in lifetime parameters that fell within the errors stated. A baseline subtraction using the 1750 cm^{-1} response was also tested, but made the fitting more challenging and the results did not improve on the treatment of the raw data. A brief comparison of these data treatments are shown in Table 3.5, with uncertainties given by standard errors.

Table 3.5: Results of Fitting and Arrhenius Analysis for Raw, D_2O and Baseline corrected kinetics

	RNA_{all}				**DNA_{all}**			
	τ_1[a] (µs)	τ_2[a] (µs)	$E_{a,m}$	$E_{a,r}$	τ_1[a] (µs)	τ_2[a] (µs)	$E_{a,m}$	$E_{a,r}$
Raw	6.9±0.1	190±13	83.3±19.6	-48.2±5.0	1.2±0.1	158±33	64.0±1.2	-45.2±5.6
-1750 cm^{-1}	6.3±0.1	199±13	82.3±22.2	-41.0±5.2	0.8±0.1	185±31	70.1±3.5	-25.1±2.5
-D_2O	6.4±0.2	192±12	86.7±22.8	-41.1±4.4	1.0±0.1	154±12	65.8±0.7	-44.5±1.9

[a)] Lifetimes for T_0 = 65 °C

Arrhenius and Eyring Analysis

In order to determine the activation energies (E_a) associated with the melting and refolding processes, the temperature dependence of the lifetimes of melting and refolding was analysed. For both RNA_{all} and DNA_{all} the temperature-dependence of the melting timescales (τ_1) was found to be in good agreement with an Arrhenius representation (Fig. 3.12 (a), above D_2O).

Therefore, the activations energies for the process could be extracted from the dynamics according to the Arrhenius equation:

$$k = Ae^{\frac{-E_a}{RT}} \qquad 3.6$$

The rate constant is k, while R is the ideal gas constant, T the temperature and A the pre-exponential factor. This yielded positive activation energies (ΔE = 83 ± 20 and 64 ± 10 $kJmol^{-1}$ for RNA and DNA hairpins respectively), as would be expected for strand melting. These values are within a factor of three compared to prior studies on double stranded DNA and RNA, though different sequences and conditions preclude direct comparison.[14,16]

To obtain details on the transition enthalpy and entropy and gain a greater insight into the varying determinants of the rates, an Eyring analysis was also conducted according to:

$$k = \frac{k_B T}{h} e^{-\frac{\Delta H^{\ddagger}}{RT}} e^{\frac{\Delta S^{\ddagger}}{R}} \qquad 3.7$$

Here the rate constant is related to the transition entropy, $\Delta S^{\ddagger}$ and enthalpy, $\Delta H^{\ddagger}$, by the Boltzmann constant, k_B, and h, Planck's constant. The full set of parameters determined by both Arrhenius and Eyring analysis can be found in Table 3.6.

Table 3.6: Results of Arrhenius and Eyring Analysis of RNA_{all} and DNA_{all}

		RNA_{all}	DNA_{all}	
Arrh	$E_{a,m}$	83.3	64.0	$kJmol^{-1}$
	$E_{a,r}$	-48.2	-45.2	$kJmol^{-1}$
Eyring	$\Delta H_m^{\ddagger}$	80.4	61.2	$kJmol^{-1}$
	$\Delta S_m^{\ddagger}$	86	46	$JK^{-1}mol^{-1}$
	$\Delta G_m^{\ddagger *}$	53.8	46.8	$kJmol^{-1}$
	$\Delta H_r^{\ddagger}$	-51.2	-48.1	$kJmol^{-1}$
	$\Delta S_r^{\ddagger}$	-324	-315	$JK^{-1}mol^{-1}$
	$\Delta G_r^{\ddagger *}$	49.1	49.5	$kJmol^{-1}$

*ΔG calculated at 37°C

In contrast to τ_1 however, the timescales of τ_2 were found to increase with increasing T_0, yielding values from 150-400 μs, from 60-80 ˚C for both RNA_{all} and DNA_{all}; consistent with a process with an apparent negative activation energy (Fig. 3.12 (a), below D_2O). As the value of T_0 decreases the timescales for RNA_{all} and DNA_{all} signal recovery decrease such that at a T_0 of 65 ˚C the hairpin dynamics match those of the solvent. We thus conclude that, at T_0 values equal to or greater than 70 ˚C the experiment is probing the native dynamics of the hairpin, which are independent of the solvent behaviour. Specifically, we conclude that hairpin refolding dynamics are being observed.

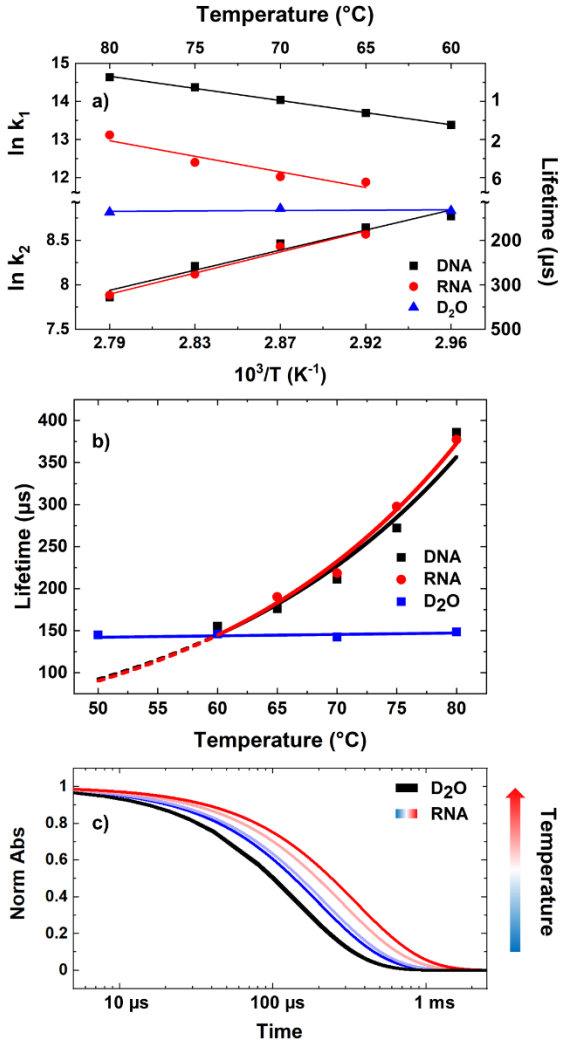

Figure 3.12: *Results of fitting T-jump spectroscopy over a range of T_0 values for RNA_{all} and DNA_{all} samples. The data is represented using a) Arrhenius plots showing the temperature dependence of the cooling rate of D_2O (blue), the melting (a, above D_2O) and refolding (a, below D_2O) of DNA_{all} (black) and RNA_{all} (red) hairpins. b) The temperature dependent lifetimes of refolding (τ_2) of DNA_{all} (black) and RNA_{all} (red) are plotted on a linear axis alongside the temperature invariant cooling timescale obtained for D_2O (blue). Arrhenius fits (lines, black and red) are shown and extrapolated for temperatures where the expected lifetimes are faster than the experimental lower limit of the D_2O cooling lifetime (dotted lines). c) Normalised τ_2 exponentials of the fit at T_0 +5 from 70-85 °C for D_2O (black) and RNA_{all} (blue-red) to illustrate the increasing hairpin refolding lifetime with temperature compared to the temperature independent D_2O cooling lifetimes.*

Using an Arrhenius analysis (Eq. 3.6) produced values of -48 $kJmol^{-1}$ ± 5 and -45 ± 6 $kJmol^{-1}$ for the activation energies of refolding of RNA_{all} and DNA_{all} respectively. These values are in agreement with negative activation energies that have been calculated for the refolding of oligonucleotides in other studies, and consistent with the direct measurement of hairpin refolding, validating the models used to obtain refolding timescales.[14,16–23] It is noteworthy that although RNA_{all} and DNA_{all} show very different melting behaviours, the refolding process is consistent between the two, an outcome which has been reported by previous studies on RNA and DNA.[24] It is important to note that the refolding process as measured is a convolution of the solvent cooling and the nucleic acid dynamics. When these timescales are well separated, as at high temperatures here, the results will mainly reflect the nucleic acid dynamics. At lower temperatures a degree of convolution may affect the absolute quantitative accuracy, but even under these conditions, a direct experimental measurement obtained for different molecules under identical experimental conditions will yield valuable comparison data.

3.4 Discussion

The marked differences in the melting timescales of the RNA_{all} and DNA_{all} hairpins, 6 vs 0.8 μs at a T_0 of 70 °C respectively, indicates that specific structural factors must influence the energetic barrier of the rate limiting step in the breaking down of the base-paired stem. While the differences between RNA and DNA at the molecular level are limited to a 2'-OH group on the ribose moiety, it is perhaps relevant that they typically adopt different helical geometries in solution. Specifically, the packing of the phosphate groups in the backbone and the stacking of the nucleotide bases are tighter in RNA than in DNA (Fig. 1.3 and

1.4 in Chapter 1).[25] X-ray structures of RNA have also determined that individual water molecules link the 2'-hydroxyl groups of adjacent nucleotides and in general have more extensive, ordered hydration of the backbone.[26] In solution, this has been shown to correlate with additional coupling between the vibrational modes of the backbone in RNA compared to DNA.[27,28] Such changes in hydration could have implications for the overall stability of the hairpin structure and its dynamics. In particular the melting process is a combination of the breakdown of base-pairing and base-stacking, but the association and dissociation of the base-stacks has been suggested to be slower than base-pairing. [29–31] This implies that the conformation and hydration dynamics that provide added stability for RNA also influence the relative stability of the base-stacking compared to DNA. The structure of some UNCG tetraloops is also known to specifically slow melting and this may have a greater effect in RNA.[11]

In contrast to melting, the refolding processes for the RNA and DNA hairpins occur on the same 200 μs timescale. The observation here of an apparent negative activation energy shows that the refolding follows a more complex mechanism than melting, which is consistent with previous studies.[14,17,22,23,29] The appearance of a rate that decreases with increased temperature has been associated with a rugged potential energy landscape such that refolding is not a simple two state transition.[21,32–34] Although there are differences between diffusive reassociation of double stranded DNA and refolding of a hairpin loop, some common elements are to be expected. To refold from a random coil the hairpin must first make an initial base pairing contact. There are a number of ways in which a base pair may be mismatched in such a way that it is in fact not possible for the stem sequence to fully base pair, and therefore these are likely to return to the random coil conformation on a number of occasions before full refolding occurs.[21] Additionally, the initial base pair must be stabilised by base

stacking to allow the full zipping of the stem.[29] Since the rate of base stacking association is relatively slow, this is likely to result in even correct initial base pairing contacts more often returning to random coil than actually resulting in the full refolding of the stem-loop. With many potential states on the energy landscape, the zipping of the hairpin represents a significant loss of entropy, which, despite also leading to a lower enthalpy, results in a free energy barrier.[14,17,21] When the temperature is increased, the rates of dissociation of base pairs and stacks is increased, making the conjunction of the two less likely and resulting in an overall slower refolding rate.[29,35] This can be confirmed by applying an Eyring rather than Arrhenius analysis, which gives access to activation enthalpy and entropy values. For the refolding process a negative enthalpy change of activation is observed that is identical to the Arrhenius activation energy (-51.2 ± 5.0 and -48.1 ± 5.6 kJ mol^{-1} for RNA and DNA respectively). This is accompanied by a negative activation entropy (-324 ± 37 and -315 ± 43 J K^{-1} mol^{-1}), such that the Gibbs free energy of activation is positive at elevated temperature, consistent with the observations. The fact that this negative activation energy and indeed the timescale of refolding is shared between both the RNA_{all} and DNA_{all} hairpin suggests that the rate limiting step for both is the same; the reforming of the base stacks to stabilise a correctly positioned initial base-pair. Furthermore, this slow rate of formation of the base stacks has a more significant impact on the refolding process than the differing features of RNA and DNA.

3.5 Conclusions

It was established that by using a thin cell, the solvent could be made to cool faster than the refolding of the tetraloop hairpins and were found to be invariant with temperature. This allowed the development of a combined T-jump followed by a T-drop experiment, making it possible to measure the native dynamics of both RNA and DNA hairpin melting and refolding for the first time. Overall we provide a direct measurement of parameters that previously had to be calculated from the melting rates and equilibrium constants. The use of IR spectroscopy for label-free detection of the conformational changes will also enable future comparisons with, for example, refolding rates obtained via microfluidic temperature drop techniques employing incorporation of fluorescent labels.[36] We have shown that the cooling timescale for a sample varies sensitively with the sample path length, meaning that this strategy can be used to tune the cooling time to a certain degree in order to match the dynamics of the analyte molecule.[5]

Comparing our results obtained on similar RNA and DNA hairpins reveals distinct melting dynamics, with RNA_{all} melting an order of magnitude more slowly than DNA_{all} (6.0 ± 0.1 μs versus 0.8 ± 0.1 μs at 70 °C), which we attribute to the influence of the helical conformation on the stabilisation of the base stacking. However, the refolding dynamics both occur on similar timescales around 200 μs and show behaviour which is consistent with a negative activation energy. These measurements are in good agreement with previous calculated values and give insights into the energy landscape that result in the similar refolding dynamics of the DNA and RNA hairpins.

3.6 References

1 M. Abdelkafi, N. Leulliot, V. Baumruk, L. Bednárová, P. Y. Turpin, A. Namane, C. Gouyette, T. Huynh-Dinh and M. Ghomi, *Biochemistry*, 1998, **37**, 7878–7884.

2 A. L. Stancik and E. B. Brauns, *J. Phys. Chem. B*, 2013, **117**, 13556–13560.

3 C. P. Howe, G. M. Greetham, B. Procacci, A. W. Parker and N. T. Hunt, *J. Phys. Chem. Lett.*, 2022, **13**, 9171–9176.

4 R. Fritzsch, G. M. Greetham, I. P. Clark, L. Minnes, M. Towrie, A. W. Parker and N. T. Hunt, *J. Phys. Chem. B*, 2019, **123**, 6188–6199.

5 G. M. Greetham, I. P. Clark, B. Young, R. Fritsch, L. Minnes, N. T. Hunt and M. Towrie, *Appl. Spectrosc.*, 2020, **74**, 720–727.

6 L. Minnes, G. M. Greetham, D. J. Shaw, I. P. Clark, R. Fritzsch, M. Towrie, A. W. Parker, A. J. Henry, R. J. Taylor and N. T. Hunt, *J. Phys. Chem. B*, 2019, **123**, 8733–8739.

7 M. Banyay, M. Sarkar and A. Gräslund, *Biophys. Chem.*, 2003, **104**, 477–488.

8 C. Lee, K. H. Park and M. Cho, *J. Chem. Phys.*, 2006, **125**, 114508.

9 C. Lee and M. Cho, *J. Chem. Phys.*, 2006, **125**, 114509.

10 C. Hyeon and D. Thirumalai, *Biophys. J.*, 2012, **102**, L11–L13.

11 D. J. Proctor, H. Ma, E. Kierzek, R. Kierzek, M. Gruebele and P. C. Bevilacqua, *Biochemistry*, 2004, **43**, 14004–14014.

12 G. Pljevaljčić, D. P. Millar and A. A. Deniz, *Biophys. J.*, 2004, **87**, 457–467.

13 G. Pljevaljčić, D. Klostermeier and D. P. Millar, *Biochemistry*, 2005, **44**,

4870–4876.

14 R. J. Menssen and A. Tokmakoff, *J. Phys. Chem. B*, 2019, **123**, 756–767.

15 B. Ashwood, N. H. C. Lewis, P. J. Sanstead and A. Tokmakoff, *J. Phys. Chem. B*, 2020, **124**, 8665–8677.

16 M. I. Wallace, L. Ying, S. Balasubramanian and D. Klenerman, *Proc. Natl. Acad. Sci. U. S. A.*, 2001, **98**, 5584–5589.

17 L. E. Revell and B. E. Williamson, *J. Chem. Educ.*, 2013, **90**, 1024–1027.

18 J. A. Wyer, M. B. Kristensen, N. C. Jones, S. V. Hoffmann and S. B. Nielsen, *Phys. Chem. Chem. Phys.*, 2014, **16**, 18827–18839.

19 T. E. Ouldridge, P. Šulc, F. Romano, J. P. K. Doye and A. A. Louis, *Nucleic Acids Res.*, 2013, **41**, 8886–8895.

20 M. C. Lin and R. B. Macgregor, *Biochemistry*, 1997, **36**, 6539–6544.

21 A. Ansari, S. V. Kuznetsov and Y. Shen, *Proc. Natl. Acad. Sci. U. S. A.*, 2001, **98**, 7771–7776.

22 P. J. Sanstead, P. Stevenson, A. Tokmakoff and A. Tokmako, *J. Am. Chem. Soc.*, 2016, **138**, 11792–11801.

23 P. J. Sanstead and A. Tokmakoff, *J. Phys. Chem. B*, 2018, **122**, 3088–3100.

24 C. H. Cheng, K. Ishii and T. Tahara, *J. Phys. Chem. B*, 2020, **124**, 10673–10681.

25 I. Anosova, E. A. Kowal, M. R. Dunn, J. C. Chaput, W. D. V. Horn and M. Egli, *Nucleic Acids Res.*, 2016, **44**, 1007–1021.

26 M. Egli, S. Portmann and N. Usman, *Biochemistry*, 1996, **35**, 8489–8494.

27 E. M. Bruening, J. Schauss, T. Siebert, B. P. Fingerhut and T. Elsaesser, *J.*

Phys. Chem. Lett., 2018, **9**, 583–587.

28 A. Kundu, J. Schauss, B. P. Fingerhut and T. Elsaesser, *J. Phys. Chem. B*, 2020, **124**, 2132–2138.

29 W. Zhang and S. J. Chen, *Biophys. J.*, 2006, **90**, 765–777.

30 K. Sarkar, K. Meister, A. Sethi and M. Gruebele, *Biophys. J.*, 2009, **97**, 1418–1427.

31 K. Sarkar, D. A. Nguyen and M. Gruebele, *RNA*, 2010, **16**, 2427–2434.

32 H. Ma, D. J. Proctor, E. Kierzek, R. Kierzek, P. C. Bevilacqua and M. Gruebele, *J. Am. Chem. Soc.*, 2006, **128**, 1523–1530.

33 M. Hairong, C. Wan, A. Wu and A. H. Zewail, *Proc. Natl. Acad. Sci. U. S. A.*, 2007, **104**, 712–716.

34 R. K. Nayak, O. B. Peersen, K. B. Hall and A. Van Orden, *J. Am. Chem. Soc.*, 2012, **134**, 2453–2456.

35 W. Zhang and S. J. Chen, *Proc. Natl. Acad. Sci. U. S. A.*, 2002, **99**, 1931–1936.

36 M. E. Polinkovsky, Y. Gambin, P. R. Banerjee, M. J. Erickstad, A. Groisman and A. A. Deniz, *Nat. Commun.*, 2014, **5**, 5737.

4. Sequence Dependent Melting and Refolding Dynamics of RNA UNCG Tetraloops Using Temperature-Jump/Drop Infrared Spectroscopy

4.1 Introduction

The work in this chapter explores the differences between the melting and refolding dynamics of RNA and DNA hairpins by investigating the sequence dependence of their conformational dynamics. By doing this, greater insight can be gained on the structural and dynamic basis for the different behaviour of DNA and RNA in biological systems despite their similar molecular composition. This chapter extends the initial study in Chapter 3 of UNCG tetraloops with all GC-containing stems by investigating three hairpins with stem sequences each featuring a single AU (RNA, Fig. 4.1) inserted at specific points in the otherwise GC stem. This was contrasted with the equivalent TNCG DNA hairpins with AT pairs introduced into the stem.

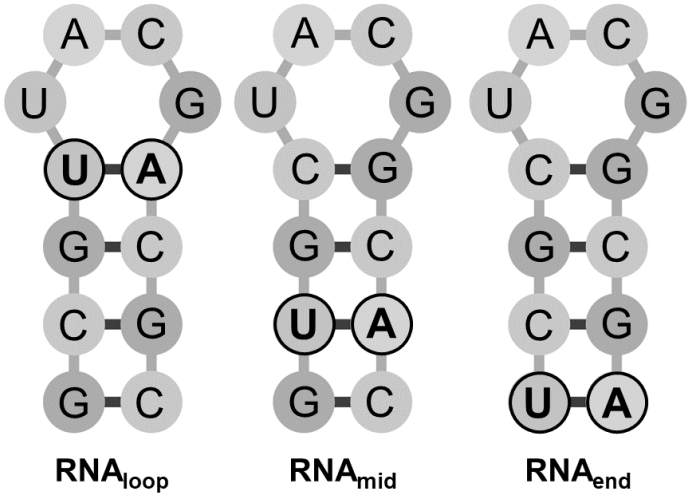

Figure 4.1: *Schematic representation of the base sequences of the RNA UACG tetraloop hairpins RNA_{loop} (left), RNA_{mid} (centre) and RNA_{end} (right) studied with T-jump/drop IR spectroscopy. The UA base pair position in the stem is highlighted in each case. The 5' end is located at the bottom left of each structure.*

To understand the details of the interactions in the hairpin stem and the mechanism of refolding we use infrared spectroscopy to probe the response of the sequences following T-jump initiation allowing clear differentiation of effects involving the main GC-rich stem and the single AU/T 'marker', providing position-specific insight into changes in stem behaviour arising from the AU/T inclusion.

The three marker positions studied in RNA hairpins were the *loop* closing base pair (Fig. 4.1, RNA_{loop}), to determine the importance of the GC closing pair upon the rate of melting and refolding dynamics of the loop; the terminal *end* pair (Fig. 4.1, RNA_{end}) to compare the difference between the two ends of the stem, and in the centre of the stem adjacent to the terminal end (Fig. 4.1, RNA_{mid}). The data for the sequence with an *all* GC stem (RNA_{all}) characterised in Chapter 3 was used as a baseline comparison. In all cases, the DNA counterparts were also studied to extend our understanding of any dynamic differences between the two molecules.

4.2 Experimental

The RNA and DNA oligomer sequences, R/DNA$_{loop}$: 5'-GCGX(XACG)ACGC-3', R/DNA$_{mid}$: 5'-GXGC(XACG)GCAC-3' and R/DNA$_{end}$: 5'-XCGC(XACG)GCGA-3' (RNA: X=U; DNA X=T; "(…)" indicates the loop position) were prepared as outlined in Chapter 2, §2.1.

FT-IR Absorption measurements of the sequences were conducted as described in Chapter 2, §2.2. Spectra were recorded over the temperature range from 20 – 80 °C through the melting with points taken at intervals of 3 °C

The temperature-jump measurements were conducted as described previously in Chapter 2, §2.3.[1–4] Chapter 3 established the use of a sample path length of 12 μm to provide ~150 μs cooling of the solvent to allow direct observation of nucleic acid refolding dynamics (see below and Chapter 3).

4.3 Results

4.3.1 Infrared Absorption Spectroscopy

Infrared absorption spectra of the eight oligonucleotide hairpin sequences (RNA$_{all/loop/mid/end}$ and DNA$_{all/loop/mid/end}$) were recorded at temperature from 20 to 80 °C and show a number of spectral changes through the melting transition (Fig. 4.2(a-h)), in the region of the mid-IR spectrum from 1550 to 1700 cm^{-1} associated with the nucleotide bases.[5] In all eight hairpins, the G_R mode at 1575 cm^{-1} (Fig. 4.2, purple panels) was observed to undergo a significant increase in intensity between 20 °C (Fig. 4.2, blue traces) and 80 °C (Fig. 4.2, red traces).

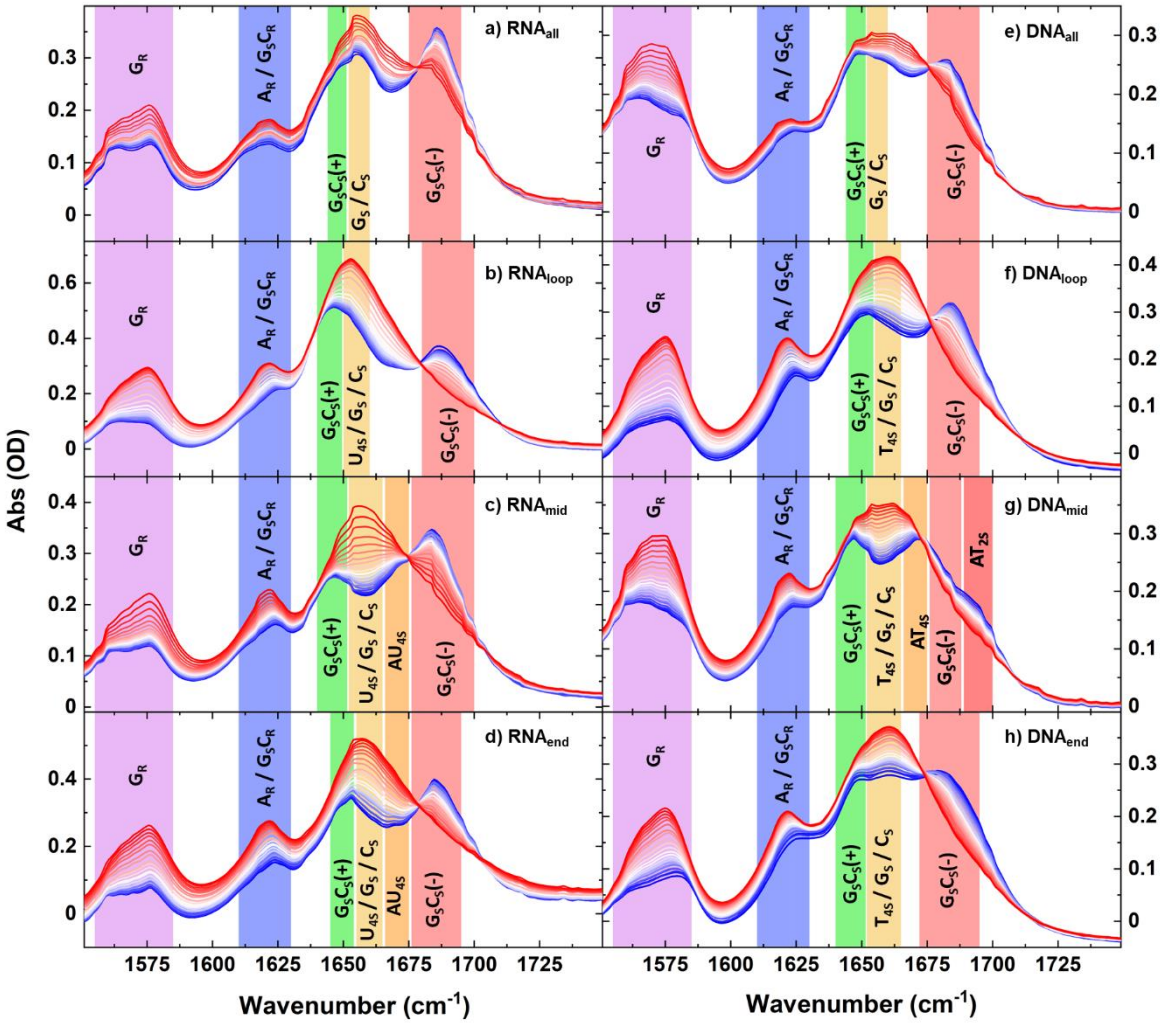

Figure 4.2: *FT-IR spectra of RNA and DNA hairpins from 20 – 80 °C (blue to red) for RNA$_{all}$, RNA$_{loop}$, RNA$_{mid}$ and RNA$_{end}$ (a-d) and DNA$_{all}$, DNA$_{loop}$, DNA$_{mid}$ and DNA$_{end}$ (e-h). Vibrational mode assignments are indicated by coloured panels: guanine ring mode, G$_R$ (1575 cm^{-1}, purple); overlapping adenine ring, A$_R$ and base paired guanine/cytosine ring mode, G$_S$C$_R$ (1620 cm^{-1}, blue); symmetric base paired guanine/cytosine mode, G$_S$C$_S$(+) (1648 cm^{-1}, green); overlapping guanine, cytosine and uracil/thymine stretching modes G$_S$, C$_S$ and U/T$_{4S}$ (1656 cm^{-1}, yellow); base paired adenine-uracil/thymine stretching mode, U/AT$_{4S}$ (orange, 1665 cm^{-1}); base paired asymmetric guanine-cytosine mode, G$_S$C$_S$(-) (1686 cm^{-1}, red) and base paired adenine and uracil/thymine stretching mode, AT$_{2S}$ (1690 cm^{-1}, dark red).*[5,9,10] *The notation used is consistent with that used in Refs 5, 9 & 10. Spectra have been solvent corrected.*

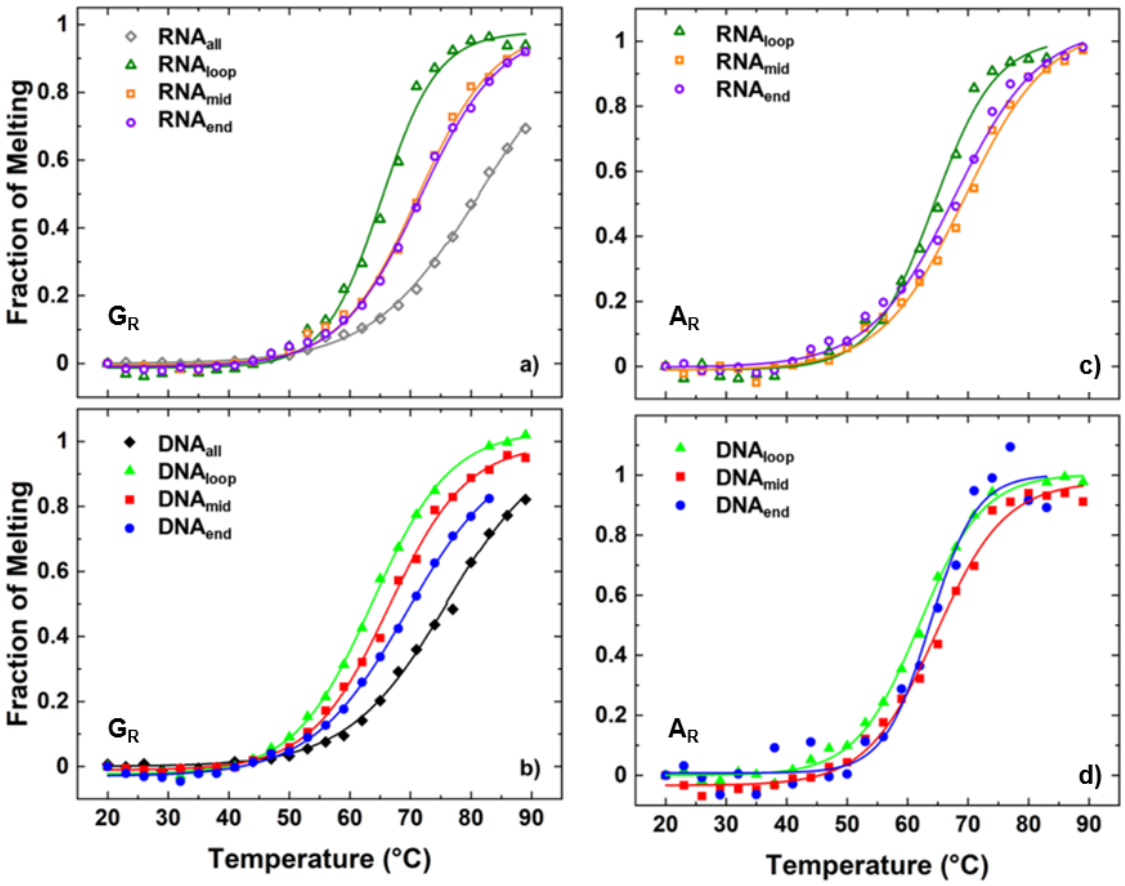

Figure 4.3: *Melting curves with respect to the G_R mode (1575 cm^{-1}) (a, b) and A_R mode at 1622 cm^{-1} (c, d) for RNA sequences (top) and DNA sequences (bottom). These were derived by plotting the intensity change of the G_R or A_R as a function of temperature. Lines show the results of fitting the data using sigmoidal functions. All datasets are normalised to their sigmoid fit maximum to show fraction of melting. All spectra were baseline corrected at 1750 cm^{-1}.A linear temperature-dependent contribution to the melting curves were extrapolated from the 20 – 40 °C range where no melting occurs and subtracted to give the data shown above.*

The melting curve of the G_R for each hairpin was used to extract the melting temperature (T_m) and equilibrium thermodynamic parameters via the Van't Hoff equation (Table 4.1). Uncertainties shown are standard errors. The Van't Hoff analysis is discussed in more detail in Chapter 3, §3.3.1.

Table 4.1: Thermodynamic parameters obtained from the G_R and A_R melting curves via Van't Hoff analysis

		RNA_{loop}		RNA_{mid}		RNA_{end}		
		G_R	A_R	G_R	A_R	G_R	A_R	
Van't Hoff	T_m	65.2 ± 0.4	64.4 ± 0.5	71.2 ± 0.5	69.4 ± 0.6	71.6 ± 0.4	67.7 ± 0.6	°C
	ΔH	203.9 ± 14.3	184.8 ± 16.5	154.4 ± 9.2	140.5 ± 10.7	148.6 ± 6.7	137.2 ± 10.7	$kJmol^{-1}$
	ΔS	603 ± 42	548 ± 49	449 ± 27	410 ± 31	431 ± 19	403 ± 31	$JK^{-1}mol^{-1}$
	ΔG[a]	17.0 ± 19	15.0 ± 22.5	15.4 ± 12.5	13.3 ± 14.4	14.9 ± 8.9	12.4 ± 14.4	$kJmol^{-1}$

		DNA_{loop}		DNA_{mid}		DNA_{end}		
		G_R	A_R	G_R	A_R	G_R	A_R	
Van't Hoff	T_m	63.7 ± 0.2	61.9 ± 0.3	66.4 ± 0.3	64.7 ± 0.5	69.3 ± 0.5	63.7 ± 0.8	°C
	ΔH	154.8 ± 3.7	178.0 ± 9.0	148.4 ± 6.4	165.9 ± 13.4	123.9 ± 6.3	188.8 ± 43.7	$kJmol^{-1}$
	ΔS	46.1 ± 11	531 ± 27	437 ± 19	491 ± 40	362 ± 18	557 ± 130	$JK^{-1}mol^{-1}$
	ΔG[a]	12.0 ± 5.0	13.3 ± 12.3	12.9 ± 8.7	13.6 ± 18.3	11.7 ± 8.4	16.1 ± 59.4	$kJmol^{-1}$

[a] All ΔG values were calculated at 37°C.

The T_m values are very similar for all six sequences with an AT or AU base pair included in the stem (loop, mid and end), with all falling within the range 68 ± 5 °C. These values are all lower than the T_m exhibited by the R/DNA_{all} sequences, which yielded T_m values of 81 and 76 °C (± 1 °C) respectively.[1] The implication is that replacing a GC base pair with an AT or AU pair results in a reduction in stability of the stem, irrespective of the position of the marker pair.

Analysing the trends in the RNA melting curves, RNA_{loop} has the lowest melting point at 65 °C and the steepest slope over the melting transition (Fig. 4.3 (a),

dark green), while the curves measured for RNA_{mid} and RNA_{end} have very similar transitions to each other and melting points of 71 and 72 °C (Fig. 4.3 (a), orange and violet). For the DNA sequences, DNA_{loop} (Fig. 4.3 (b), green) has a melting point of 64 °C, compared to 66 °C for DNA_{mid} (Fig. 4.3 (b), red) and 69 °C for DNA_{end} (Fig. 4.3 (b), blue). The A_R signal for DNA_{all} was less than half of the other sequences, resulting in worse signal to noise and a higher degree of uncertainty in the sigmoid fit.

In the sequences $RNA_{loop/mid/end}$ and their DNA equivalents, the A_R band at 1625 cm^{-1} was observed to undergo a rise in intensity with increasing temperature (Fig. 4.2, blue panels). [5,10] The temperature dependence of the band intensity was also found to be sigmoidal in nature (Fig. 4.3, (c, d). The amplitude change is assigned to loss of base-stacking and pairing of the single AU (RNA) and AT (DNA) base pairs placed in the stems of the $R/DNA_{loop/mid/end}$ hairpins respectively in a similar manner to that of the G_R mode.

As only a single base pair contributes to the intensity change of the A_R band upon stem melting, the increase in magnitude of the A_R band was found to be much smaller than that of the G_R band. Nevertheless, this mode serves as a proxy for the behaviour of the AT/AU base pair in each hairpin and will be used as such in T-jump experiments below. It is however important to be aware of overlapping contributions from a number of different modes in this region which do not affect the G_R band.[5,10]

Site specific equilibrium thermodynamic parameters were determined (Table 4.1) via a Van't Hoff analysis of the A_R-derived melting curves.

RNA Sequence-Dependent Variations

The G_R and A_R modes allow assignment of the changes in the IR absorption spectra of the nucleic acid sequences as the temperature is increased to melting of the base paired stem. By association, we assume concomitant loss of the hairpin structure. It is however also instructive to consider sequence specific variations in the IR absorption spectra of the hairpins in the more complex spectral region between 1630 and 1700 cm^{-1}.

Considering the spectra of the RNA hairpins ($RNA_{all/loop/mid/end}$) at 20 °C (Fig. 4.2 (a-d), blue traces) it can be seen that two prominent bands near 1648 cm^{-1} and 1686 cm^{-1} appear in all four spectra, which are attributable to the (G_SC_S(+)) and (G_SC_S(-)) modes respectively (Fig 4.2: green and red panels).[5,10] As the temperature is increased, G_SC_S(+) shifts slightly to higher wavenumber and increases in intensity as the strands melt while the G_SC_S(-) undergoes a large loss in intensity. More formally, the base paired modes (G_SC_S(+/-)) are replaced by the uncoupled G_S and C_S modes of the unpaired bases (Fig. 4.2, yellow), which results in a large, broad peak near 1660 cm^{-1} in the spectra of all four sequences obtained at 80 °C (Fig. 4.2 (a-d), red traces).

Comparing the IR absorption spectra of all of the RNA hairpins at 80 °C (Fig. 4.2 (a-d), red traces, Fig. 4.4, dashed line) shows that they are broadly similar, as would be expected when no secondary structure exists, though a small amount of the G_sC_S(-) mode is still visible in the RNA_{all} spectrum at 80 °C (Fig. 4.2 (a)), as a result of its slightly higher T_m value than the other hairpins. At 20 °C (Fig. 4.2, blue traces and Fig. 4.4, solid lines), other differences between the spectra of RNA_{all} and the $RNA_{loop/mid/end}$ sequences become clear, which we infer arise from the inclusion of a single AU base pair at different points in the stems of $RNA_{loop/mid/end}$.

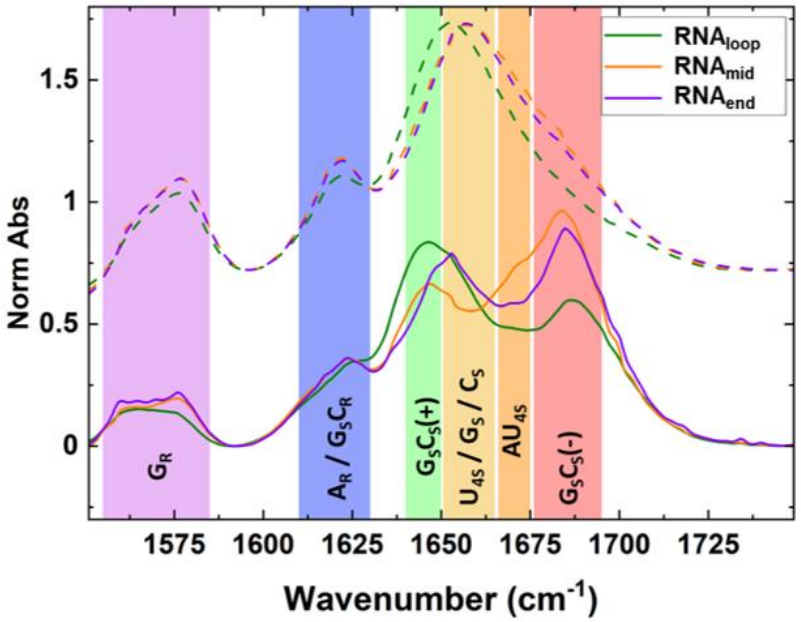

Figure 4.4: *Normalised IR absorption spectra obtained at 20 °C (solid lines) and 80 °C (dashed lines) for RNA_{loop} (dark green), RNA_{mid} (orange), and RNA_{end} (violet). Prominent vibrational modes are indicated by coloured panels (see main text). Spectra have been solvent and baseline corrected. All spectra are normalised to the 80 °C spectrum at 1656 cm^{-1} for the respective sequence*.

In the case of RNA_{mid} where the AU pair is located near the centre of the GC-rich stem (Fig. 4.1) and so would be expected to cause minimal perturbation to the double stranded structure, a shoulder is observed on the low frequency side of the G_SC_S(-) band near 1675 cm^{-1} (Fig. 4.2 (c), orange panel). We assign this band to the AU_{4S} stretching mode of the AU base pair, which is mainly due to the stretching vibration of the 4-position C=O bond in the W-C base paired uracil base.[9,10,15] The observation of the AU_{4S} mode is consistent with strong base pairing of the AU in the stem of RNA_{mid} as would be expected given its position. Upon stem melting (Fig. 4.2 (c), red traces), the AU_{4S} band of RNA_{mid} is replaced by the U_{4S} mode of the unpaired uracil at 1660 cm^{-1} (yellow panel). This band overlaps with both the G_S and C_S modes, leading to a noticeably larger 1660 cm^{-1} band (Fig. 4.2 (c), yellow panel) for RNA_{mid} in comparison to RNA_{all} (Fig. 4.2 (a)).

The spectrum of RNA_{end} at 20 °C (Fig. 4.2 (d), blue) is similar to that of RNA_{mid} (Fig. 4.2 (c)). Although the AU_{4S} shoulder is not clearly visible, the G_SC_S(-) band is slightly broadened on the low frequency side in comparison to that of RNA_{all}. This suggests that in RNA_{end} the A and U modes are not as strongly coupled as in RNA_{mid}, but that the AU is still in a base paired configuration.

Conversely, the 20 °C spectrum of RNA_{loop} (Fig. 4.2 (b), blue traces) is very different to those of $RNA_{all/mid/end}$ with the relative intensities of the G_SC_S(+/-) bands reversed. Indeed, the G_SC_S(+) mode (Fig. 4.2(b), green panel) is both higher in intensity and broadened relative to those of $RNA_{all/mid/end}$ and the general shape of the spectrum of RNA_{loop} at 20 °C is more similar to the spectra of $RNA_{mid/end}$ at intermediate temperatures between 20 and 80 °C. We attribute this to a greater contribution to the 20 °C RNA_{loop} spectrum from the non-W-C U_{4S} mode. While such an observation could indicate some degree of fraying of the pair of bases at the closing point of the tetraloop, we note that a significant rise in intensity of the A_R mode (Fig. 4.2, blue panel) with temperature is still observed for RNA_{loop}. As the increase in A_R mode intensity is associated with a loss of base-stacking during the melting process, this suggests that the AU base stacking interactions are still in place at low temperature. Thus, while there appears to be some disruption to the base pairing in the neck of the RNA_{loop} hairpin, which is not seen in the $RNA_{mid/end}$ sequences, we do not attribute this to full base pair fraying.

DNA Sequence-Dependent Variations

Applying a similar analysis to the IR absorption spectra of the four DNA sequences (Fig. 4.2 (e-h)) shows that, like their RNA counterparts, these too share very similar features at 80 °C although the slightly different position of T_{4S} relative to U_{4S} gives the peak near 1660 cm^{-1} a more symmetric profile in DNA than was observed in RNA (Fig. 4.2 red spectra, Fig 4.5, dashed lines). In the DNA hairpins, the G_SC_S(-) peaks are also generally less prominent than those of RNA.

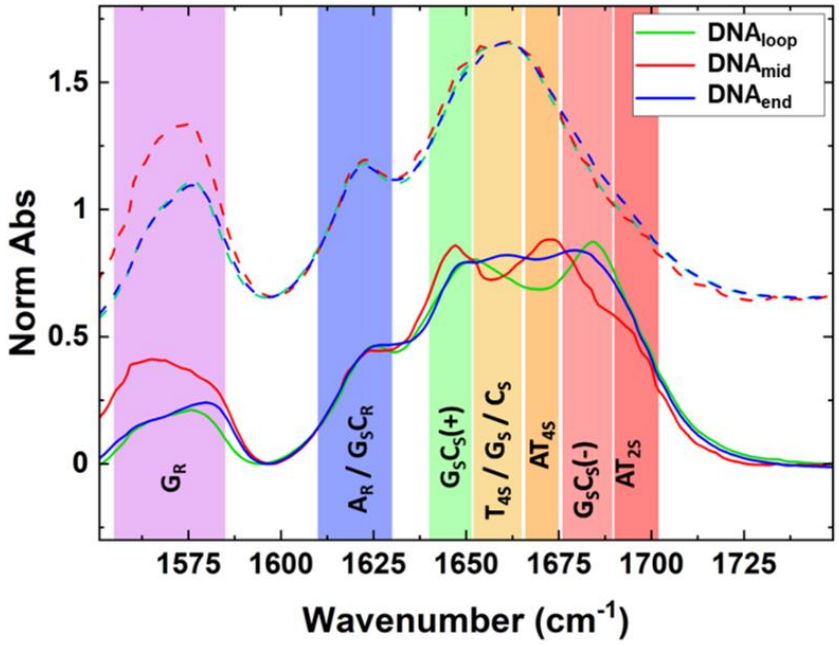

Figure 4.5: *Normalised IR absorption spectra obtained at 20 °C (solid lines) and 80 °C (dashed lines) for DNA_{loop} (green), DNA_{mid} (red) and DNA_{end} (blue). Prominent vibrational modes are indicated by coloured panels (see main text). Spectra have been solvent and baseline corrected. All spectra are normalised to the 80 °C spectrum at 1656 cm^{-1} for the respective sequence.*

Once again variations in the 20 °C spectra (Fig. 4.2 (e-h), blue) provide insight into possible sequence dependent differences in structural configurations caused by the positioning of the single AT base pair (Fig. 4.2 (e-h) blue, Fig 4.5, solid lines). At 20 °C, the spectrum of DNA_{mid} (Fig. 4.2 (g), blue trace), with the AT pair in the centre of the stem, contains contributions from the AT_{4S} (orange panel) and AT_{2S} (dark red panel) modes of the AT base pair. These bands occur to high and low frequency relative to the G_SC_S(-) mode (red panel) and so significantly alter the lineshape and peak positions in the region from 1670 -1700 cm^{-1}. The presence of peaks deriving from the base paired configuration of A and T indicate relatively strong AT base pairing in this sequence, consistent with the sharp rise of the A_R band alongside the G_R band (Fig. 4.2 (g) blue and purple panels) with increasing temperature.

In the case of the DNA_{loop} hairpin, the G_SC_S(+) and G_SC_S(-) modes (Fig. 4.2 (f), green and red panels) are of similar intensity to one another at 20 °C (blue). In this case, the AT_{4S} and AT_{2S} peaks observed in DNA_{mid} are not as clear, with AT_{2S} appearing as a weak high frequency shoulder to G_SC_S(-), and there is a similarly weak contribution from the T_{4S} mode (Fig. 4.2 (f), yellow panel) which is manifest as a broadening of the G_SC_S(+) band on its high frequency side. Together, these observations are indicative of slightly reduced integrity of the base pairing near the neck of the tetraloop. The 20 °C spectra of DNA_{all} and DNA_{loop} are however very similar, suggesting minimal change due to the AT pair. As with the RNA_{loop} sequence, a rise in intensity with temperature was observed for the A_R mode, consistent with the degree of disruption at the base of the loop being less than full fraying.

Finally, the 20 °C spectrum of DNA_{end} (Fig. 4.2 (h), blue) shows that the G_SC_S(+) and G_SC_S(-) peaks form a plateau with the T_{4S} peak appearing distinctly between them (Fig. 4.2 (h), yellow panel, Fig. 4.5, blue traces). This is a clear indicator of

a weak terminal AT pair that is fraying to a significant degree, while the GC bases remain paired. Indeed, the absolute change in amplitude of the A_R mode of DNA_{end} between 20-80 °C was found to be less than half that of the other five sequences, consistent with a much-reduced loss of AT base pairing as the stem melted, as would be anticipated if the bases in the terminal position were already frayed. This is a notable difference between the RNA and DNA sequences, where no evidence of a frayed terminal AU was observed for RNA_{end}.

4.3.2 Temperature-Jump/Drop IR Spectroscopy

Temperature-jump/drop IR spectroscopy was conducted on the $RNA_{loop/mid/end}$ (Fig. 4.1) sequences and their DNA analogues spanning a range of starting temperatures (T_0) up to the point where the T-jump-induced temperature rise crossed their respective melting transitions (T_m). This approach enabled insight into the influence of stem base sequence on the temperature dependent melting and refolding dynamics of the hairpins.

The results of T-jump/drop experiments on the R/DNA_{all} sequences discussed in Chapter 3,[1] determined that the T-jump pulse heats the solvent quickly (ns), initiating melting of the hairpin stem, with RNA_{all} found to melt an order of magnitude more slowly than DNA_{all} (6 ± 0.1 μs versus 0.8 ± 0.1 μs at 70 °C). Solvent cooling was shown to occur on timescales of < 140 μs, more quickly than hairpin refolding, allowing the refolding times of RNA_{all} and DNA_{all} to be measured, with both occurring on similar timescales of ~200 μs at 70 °C.[1]

The T-jump spectra for $RNA_{loop/mid/end}$, which all feature an AU base pair in the stem, are shown in Fig. 4.6 (a-c) for a starting temperature of T_m – 5 °C, with the equivalent data for the DNA sequences in Fig. 4.6 (d-f). In all cases, the spectra are presented as pump-on – pump-off absorbance difference spectra and show

the changes in the spectra occurring over timescales from T-jump initiation to the sequence dependent maximum signal, which was observed at T-jump-probe delay times of ~6-20 μs (Fig. 4.6).

For all RNA and DNA sequences, a sharp rise in intensity of the G_R (Fig. 4.6, purple panels) and A_R bands (Fig. 4.6, blue panels) was observed following the T-jump, establishing that stem melting is occurring. At longer T-jump-probe delay times, the peaks reduced in intensity towards the baseline. During this period, the spectral changes observed during the melting phase were reversed (See Appendix, Fig. A4.1) and, by analogy with data for R/DNA_{all}, we assign this to refolding of the hairpin stem as the sample cools following the T-jump.[1]

The melting and refolding dynamics for each the $R/DNA_{loop/mid/end}$ sequences were obtained by fitting the time dependent amplitudes of the G_R and A_R modes to triple exponential functions (Figs. 4.7 & A4.2-4). As established in Chapter 3,[1] the two major exponential time constants indicate the lifetimes of melting (τ_1) and refolding (τ_2). The third exponential term was generally of low amplitude (<25%) with lifetimes much longer than τ_2.

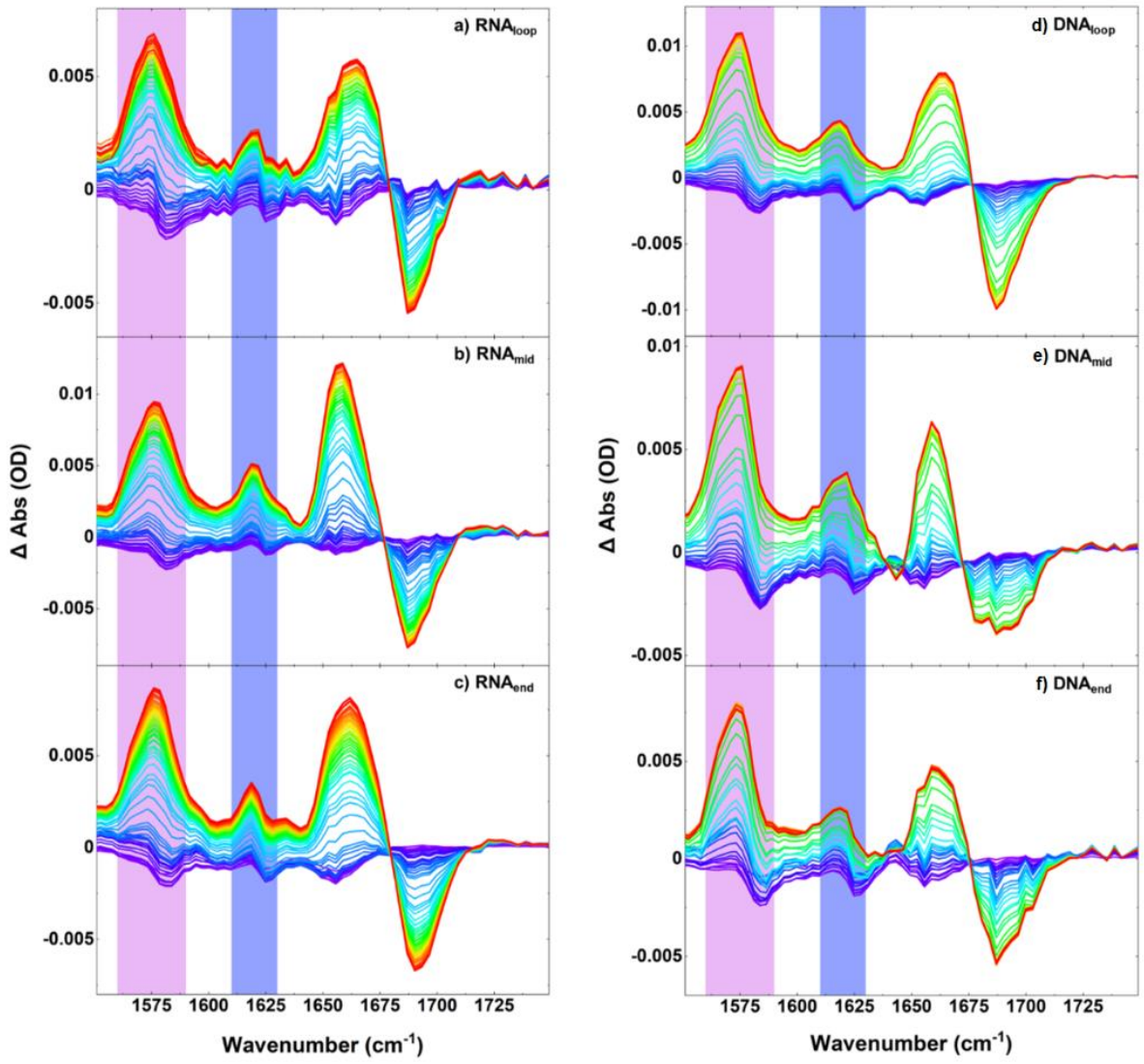

Figure 4.6: *T-jump spectra for a) RNA_{loop}, b) RNA_{mid}, c) RNA_{end} and d) DNA_{loop}, e) DNA_{mid} and f) DNA_{end}, showing the response of the hairpins from 1 ns (blue spectra) to the maximum signal obtained at delay times of ~6-20 µs (red spectra). Peak times vary between RNA and DNA and by sequence. Spectra were obtained at T_0 values of T_m – 5 °C. T-jump spectra are shown as a pump-on – pump-off difference spectra. The G_R mode at 1575 cm^{-1} and the A_R mode at 1620 cm^{-1} have been highlighted in purple (G_R) and blue (A_R). The signals observed at very early T-jump-probe delay times in both datasets (blue traces) are due to fast hydrogen bonding rearrangement.[1] Spectra have been baseline corrected for visual clarity.*

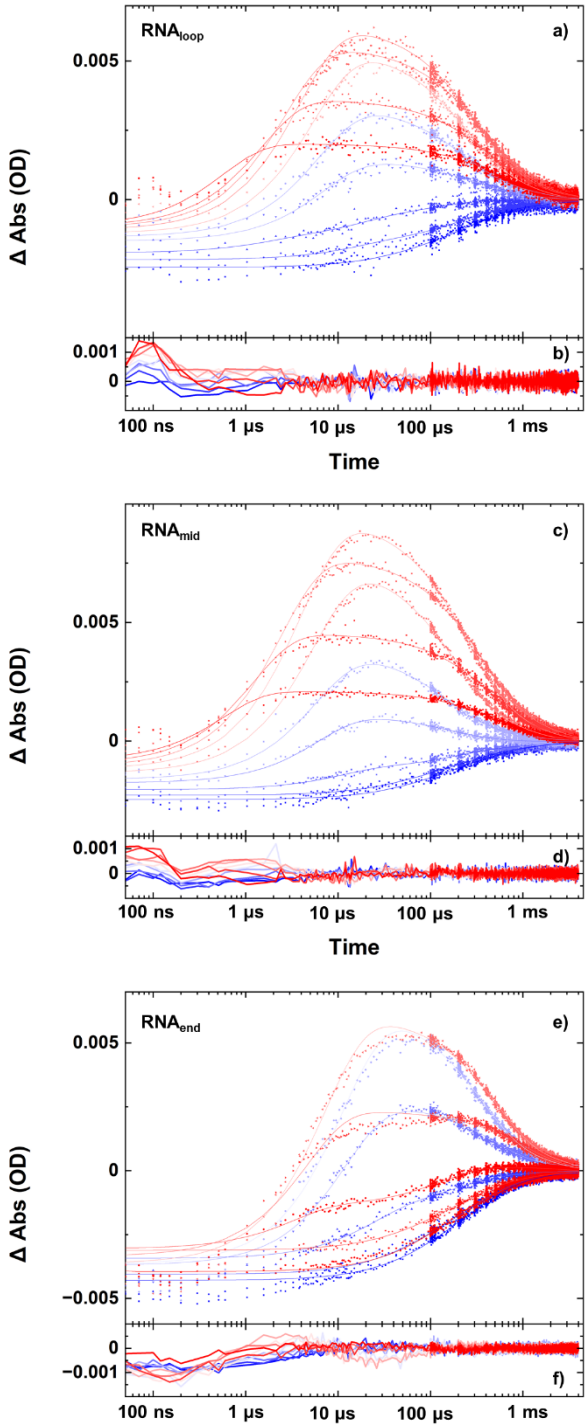

***Figure 4.7**: T-jump/drop dynamics showing temperature- and time-dependences of the G_R band of the RNA sequences a) RNA_{loop}, c) RNA_{mid} and e) RNA_{end}. Data (dots) are shown from a T_0 of 20 °C to 80 °C (blue-red) along with the results of fitting to a triple exponential function (lines, see text). Residuals following the fitting process are shown in panels (b, d, f).*

G_R Melting Dynamics

The τ_1 values obtained from fitting the time dependent amplitudes of the G_R modes of the R/$DNA_{loop/mid/end}$ sequences are shown in Fig. 4.8 where it can be seen that all follow an Arrhenius temperature profile with a positive activation energy, consistent with the work presented in Chapter 3.[1] The D/RNA_{all} data is also provided for comparison purposes. Activation energies obtained from the Arrhenius analysis are shown in Table 4.2.

The lifetimes of melting with respect to the G_R mode represent an average over all of GC base pairs in the stem, though the link between G_R mode intensity and base stacking suggests that it will provide a good measure of the overall state of the hairpin stem at a given time after the T-jump.

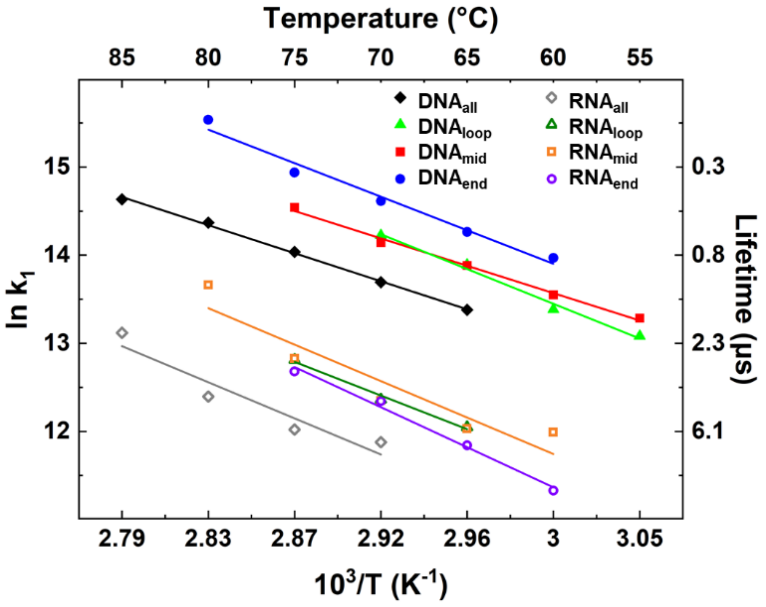

Figure 4.8: *Arrhenius analysis of the temperature dependent timescales of melting (τ_1) determined via the G_R band for RNA (open symbols) and DNA (filled symbols) sequences. The data is shown over the temperature range where the maximum change in intensity of the G_R band was greater than 20 % of the largest signal observed. Temperatures quoted are T_0 + 5 °C, the average temperature over the T-jump.*

Table 4.2: Results of Fitting and Arrhenius analysis for G_R and A_R modes

	RNA_{loop}		RNA_{mid}		RNA_{end}		
	G_R	A_R	G_R	A_R	G_R	A_R	
τ_1[a]	4.3 ± 0.4	4.3 ± 0.4	4.3 ± 0.3	4.5 ± 0.1	4.4 ± 0.3	7.6 ± 2.0	µs
τ_2[a]	299 ± 23	-	240 ± 30	-	746 ± 36	-	µs
Arr. $E_{a,m}$	74.6 ± 8.2	67.9 ± 7.0	80.4 ± 17.6	83.1 ± 20.2	70.4 ± 1.8	61.1 ± 5.8	$kJmol^{-1}$
Arr. $E_{a,r}$	-39.0 ± 0.8	-	-54.9 ± 4.3	-	-74.8 ± 0.4	-	$kJmol^{-1}$

	DNA_{loop}		DNA_{mid}		DNA_{end}		
	G_R	A_R	G_R	A_R	G_R	A_R	
τ_1[a]	0.7 ± 0.1	2.0 ± 0.1	0.7 ± 0.1	1.4 ± 0.3	0.5 ± 0.1	0.4 ± 0.1	µs
τ_2[a]	435 ± 31	-	253 ± 11	-	207 ± 8	-	µs
Arr. $E_{a,m}$	73.6 ± 6.3	25.2 ± 10.5	58.9 ± 2.9	32.2 ± 2.8	74.2 ± 6.9	54.4 ± 4.7	$kJmol^{-1}$
Arr. $E_{a,r}$	-43.2 ± 1.9	-	-41.6 ± 4.5	-	-24.0 ± 2.6	-	$kJmol^{-1}$

[a] Dynamic parameters (τ_1 and τ_2) are quoted at a temperature of 70 °C. Refolding parameters not extracted from A_R kinetics

Comparing the behaviour of all sequences at a given temperature allows discussion of relative dynamic timescales free from Arrhenius-related effects on rates. Considering the RNA sequences, the melting dynamics determined for RNA_{loop}, RNA_{mid} and RNA_{end} were consistent with one another, yielding lifetimes of 4.3 ± 0.4, 4.3 ± 0.3 and 4.4 ± 0.3 µs respectively (Fig. 4.8; dark green, orange, violet) at a T_0 + 5 °C of 70 °C. As all sequences studied here show a similar T_m, using a value of 70 °C probes a similar part of the melting curve for all sequences, as well as allowing comparisons with R/DNA_{all}. A value of T_0 + 5 °C is used to reflect the fact that the temperature jump was ~ 10 °C on average across the sample. Henceforth, all temperatures quoted will be T_0 + 5 °C unless stated.

The melting timescales of ~4.3 µs for RNA_{loop}, RNA_{mid} and RNA_{end} compare with a lifetime of melting of 6.9 ± 0.1 µs for RNA_{all} at the same temperature (Fig 5, grey). It is thus clear that inclusion of an AU base pair destabilises the hairpin stems of $RNA_{loop/mid/end}$ relative to that of the RNA_{all} sequence, while the similarities of the G_R-derived melting timescales for $RNA_{loop/mid/end}$ suggest that

the proportion of AU to GC content exerts a greater influence on the melting dynamics than the specific sequence. This is also consistent with the similar reductions in T_m observed for $RNA_{loop/mid/end}$ relative to RNA_{all} noted above.

For the DNA hairpins the general pattern observed for the RNA samples of inclusion of an AT base pair leading to a shorter hairpin melting time was also observed, with DNA_{loop} and DNA_{mid} producing effectively identical τ_1 values of 0.7 ± 0.1 μs, while DNA_{end} melted on a slightly shorter timescale at 0.5 ± 0.1 μs (at 70 °C). These compare with a value of 1.1 ± 0.1 μs for DNA_{all}. Also noticeable is that the RNA sequences all melt with considerably longer timescales than their DNA counterparts, consistent with the dynamics of RNA_{all} and DNA_{all} discussed in Chapter 3.[1] The slightly shorter melting timescale for DNA_{end} relative to DNA_{loop} and DNA_{mid} could be associated with the observation of end fraying of the terminal base pair via IR absorption spectroscopy.

A_R Melting Dynamics

Unlike the information provided by the G_R band, the dynamic information derived from the A_R band provides site specific insight, by virtue of there being only one paired adenine in each stem. By examining the dynamics of this band directly, a clearer understanding of how AU and AT inclusions at different points affects the stem can be established.

The lifetimes of melting determined for the A_R modes (Fig. 4.9) of RNA_{loop} and RNA_{mid}, 4.3 ± 0.4 μs and 4.5 ± 0.1 μs (70 °C) were indistinguishable both from one another and from the τ_1 values obtained for these sequences from the G_R mode. This indicates that the AU base pair is behaving as an integral component of the stem in these two cases. By contrast, the A_R melting timescale for RNA_{end} was found to be much longer, yielding a τ_1 of 7.6 ± 2.0 μs at the same temperature. This implies that, rather than reflecting the dynamics of the main

stem, the terminal base pair retains some degree of H-bonding or base stacking after the main part of the stem has dissociated, perhaps indicating the formation of a bubble-type structure prior to fully melting. Such a scenario would be consistent with IR absorption data suggesting that the terminal AU base pair is not prone to end fraying in RNA_{end}.

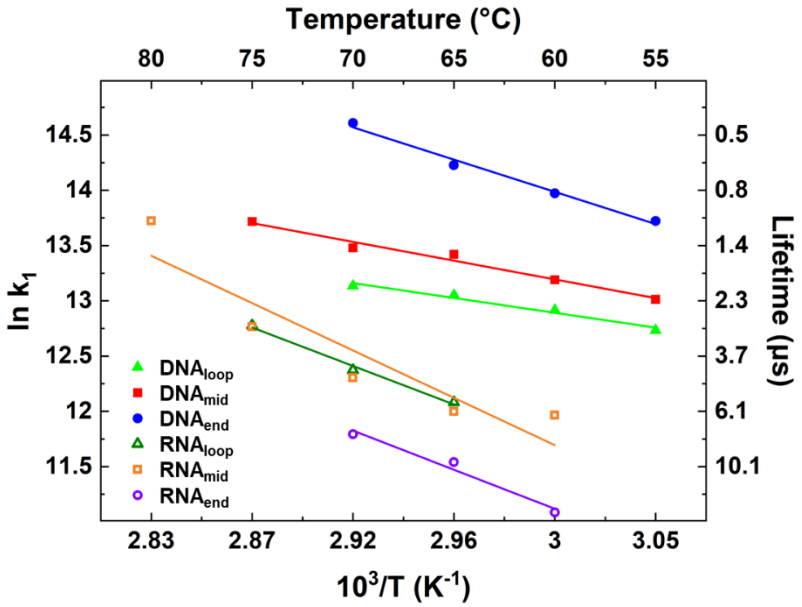

Figure 4.9: *Arrhenius analysis of the temperature dependent timescales of melting (τ_1) derived from the A_R band of RNA (open symbols) and DNA (filled symbols) sequences. The data is shown over the temperature range where the maximum change in intensity of the A_R band was greater than 20 % of the largest signal observed. Temperatures quoted are T_0 + 5 °C, the average temperature over the T-jump.*

In the case of the DNA hairpins, the A_R modes of DNA_{loop} and DNA_{mid} returned similar melting lifetimes of 2.0 ± 0.1 µs and 1.4 ± 0.3 µs respectively at 70 °C, slightly longer than the values obtained from the G_R mode for the same sequences. Conversely, the melting timescale of the A_R mode of DNA_{end} yielded a value of 0.4 ± 0.1 µs, shorter than the melting timescale for the G_R mode and consistent with DNA_{end} having a frayed terminal base pair that destabilises the stem relative to the other sequences. It is noteworthy that this melting

behaviour, where DNA_{end} shows a faster melting time than the rest of the stem, is opposite to that seen for the RNA_{end} sequence.

Refolding Dynamics

It has been established previously in Chapter 3 that the cooling dynamics of the solvent following the T-jump are temperature independent (Fig. 4.10, cyan),[1] and that the RNA_{all} and DNA_{all} hairpins were found to refold more slowly than the solvent cooling timescale of 140 μs. Furthermore, the hairpins were found to take longer to refold as the temperature was increased, which was manifest as a positive slope in the Arrhenius plot with apparent negative activation energies. These dynamics were assigned to a complex refolding landscape featuring a number of transient intermediate contacts between bases in the stem that did not lead directly to a fully base paired and stacked stem.[1,7,8,16–22] One of the consequences of this for T-drop spectroscopy is that at higher temperatures the separation between refolding timescales and solvent cooling is greatest, minimising the effects of convolution of cooling and refolding pathways.[1]

For the sequences with AT or AU base pair inclusions studied here, the refolding dynamics are reported based on only the G_R mode. This is because the G_R-derived signals were significantly larger than the A_R band, making fits of τ_2 and τ_3 more reliable.

Figure 4.10 (a) shows the refolding timescales (τ_2) obtained for all eight R/DNA sequences alongside the solvent cooling time (cyan). Anti-Arrhenius behaviour is clear for all R/DNA sequences, indicative of a complex multi-step process as previously observed for RNA_{all} and DNA_{all} (see Chapter 3). Comparing the refolding dynamics (Fig. 4.10 (b)) of the individual sequences at 70 °C yields refolding times for RNA_{mid} and RNA_{loop} as 240 ± 30 and 299 ± 23 μs, respectively.

By contrast, the value obtained for RNA_{end} was 746 ± 36 µs indicating that inserting an AU base pair at the terminal position significantly inhibits the refolding of the RNA hairpin.

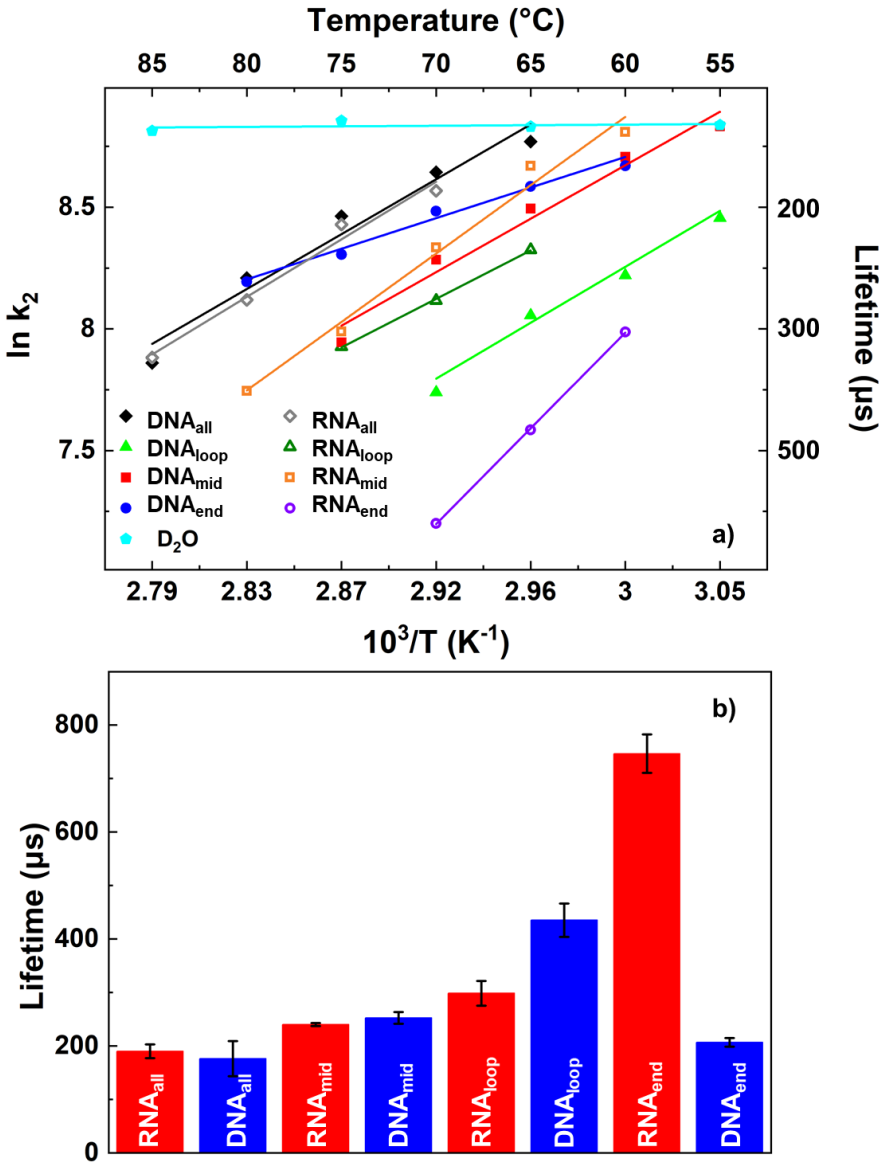

Figure 4.10: *Temperature dependence of the lifetimes of refolding (τ_2) derived from the G_R band of the RNA and DNA sequences. a) Arrhenius analysis for RNA (open symbols) and DNA (filled symbols) sequences. The D_2O solvent cooling dynamics, which are temperature-independent, are shown for comparison (cyan). Temperatures shown are T_0 + 5 °C. b) Bar graph showing a comparison of the lifetimes of refolding for all hairpins of RNA (red) and DNA (blue) at a T_0 + 5 °C of 70 °C.*

For the DNA hairpins, the trend was reversed; DNA_{end} and DNA_{mid} produced similar refolding timescales of 207 ± 8 and 253 ± 11 μs, while DNA_{loop} produced the longest τ_2 value of 435 ± 31 μs (all at 70 °C). In this case, the AT base pair near the closing point of the loop appears to perturb the refolding process most dramatically.

It was established in Chapter 3 that for the R/DNA_{all} sequences although the melting rates were significantly longer for RNA compared to DNA their refolding timescales were very similar, with values of 190 ± 13 and 176 ± 33 μs being observed respectively.[1] Comparing these values to the data for the sequences with an AT or AU inclusion measured here shows that adding an AT/AU pair results in a generally longer refolding timescale. This suggests that refolding takes place more slowly, a fact consistent with both the reduction in T_m and the faster melting rates which accompany replacement of a GC base pair with an AT/AU pair, all of which indicate a less stable stem-loop structure. It is also notable that the refolding times of RNA_{mid}, RNA_{loop}, DNA_{end} and DNA_{mid} are all comparable, being on the order of ~250 μs, as would be expected based on previous work.[1] This leaves DNA_{loop} and RNA_{end} as clear outliers and shows that placing AU/AT pairs at these positions significantly inhibits the refolding pathway in each case.

It is interesting to compare the values obtained from the Arrhenius analysis with the trends derived directly from the dynamic timescales. Focusing on the G_R-derived values shows that the activation energies for melting are all broadly similar. The values for refolding show a greater variation. In particular, the DNA_{end} and RNA_{loop} sequences (see Table 4.2) show a small, negative activation energy that differs from the other sequences. It must however be noted that Arrhenius activation energy parameters are derived from the temperature

dependence of the dynamic timescales and are enthalpic values only, thereby ignoring entropic effects, which could be substantial for mixed base sequences.

4.4 Discussion

To understand the processes occurring and relate the similarities and differences in dynamics between analogous RNA and DNA hairpins to their structures it is important to draw on all the evidence from the IR absorption spectra and T-jump/drop dynamics from the perspective of both the G_R and A_R bands.

4.4.1 General Observations

The longer melting timescale observed for the RNA_{all} hairpin compared to DNA_{all} has been attributed to differences in stem stability arising from the differences in base stacking between DNA and RNA, which can also be seen in the fact that they adopt differing helical structures, with RNA favouring the A-form and DNA the B-form.[23] These structural factors also apply to the hairpins with varied stem sequences, based on T-jump data from the G_R mode, which acts as a reporter on overall stem dynamics. The RNA sequences were still found to melt more slowly than their DNA counterparts following site specific inclusions of AU or AT base pairs into the stem. For both RNA and DNA, it was found that replacing a GC pair with an AU or AT led to a generally shorter melting timescale for the G_R mode, suggesting a relative destabilisation of the stem-loop structure. This is in line with expectation based on the stronger pairing of GC bases compared to AU/AT because the former features three H-bonds, and the latter only two.

In terms of the refolding dynamics, all of the sequences with an AU/AT inclusion exhibited longer refolding timescales than the sequence with an all-GC stem (see Table 4.2). The previous observation that RNA and DNA hairpins refolded on

similar timescales was generally found to hold, but with two clear exceptions and we focus on these in more detail below. On the whole the refolding process was found to be much slower than that of melting and can be attributed to the existence of more complex free energy profiles, with many more competing conformations and configurations on the potential energy surface to explore in comparison to strand melting. This mechanism is responsible for the anti-Arrhenius behaviour and the apparent negative activation energies observed as, when considering the multiple possible intermediates, refolding is not a simple two state process. [1,7,24–26,8,16–22] Rather, we hypothesise that it is limited by the formation of the initial, correctly positioned, base pair and first stabilising base stack. [24]

4.4.2 RNA tetraloops

Considering the G_R-derived data, it is clear that the melting timescales of $RNA_{loop/mid/end}$ are very similar and shorter than RNA_{all}. When examining the behaviour of the A_R modes for the individual structures, the site-specific labels of RNA_{loop} and RNA_{mid} track the G_R dynamics closely, enabling one to establish that they melt on comparable timescales. Such similar melting dynamics are consistent with a model where the conformation of the AU pair is governed by the rest of the GC stem. By contrast, the melting timescale of the A_R mode of RNA_{end} was found to be significantly slower than that of the other sequences, and slower than the GC melting behaviour for the same strand.

From the IR spectroscopy data, the fact that the IR absorption spectrum of RNA_{loop} at 20 °C was similar in form to the spectra of $RNA_{mid/end}$ at intermediate temperatures between 20 and 80 °C provided some evidence for a disrupted base pair at the closing point of the tetraloop, though the dynamic measurements suggest that this does not affect the melting timescales. Taking

the evidence together, the implication is of an RNA hairpin which melts from the neck of the loop towards the end of the four base pair stem.

In terms of refolding dynamics, RNA_{loop} and RNA_{mid} were again found to behave similarly, with comparable melting timescales that were slightly longer than those of the RNA_{all} sequence, consistent with a reduced driving force for reassociation. Once again, RNA_{end} was an outlier with the refolding time of this sequence being most affected by base pair substitution and much longer than the other sequences. This shows that having an AU rather than GC at the end of the stem inhibits refolding. Given that this base pair was also found to exhibit a long melting lifetime, these two pieces of evidence could imply a significant degree of flexibility in the AU base pairing geometry whereby it can accommodate a significant perturbation of its structure without breaking the H-bonds, but that these perturbed structures, when formed as transient steps on the refolding pathway, do not lead to the reformation of the stem. Overall, the data are consistent with a mechanism for RNA melting that occurs from the loop but refolds from the terminal end.

4.4.3 DNA tetraloops

Considering the G_R-derived data for the DNA sequences with an AT inclusion we observe significant differences in the spectroscopy and dynamics of the sequences, and by extension the unfolding and folding mechanisms compared to RNA. For example, DNA_{end} melted on a shorter timescale than DNA_{loop} and DNA_{mid}, which were comparable. Using site-specific data from the A_R mode emphasises this trend, with the A_R mode melting on timescales of 2.0, 1.4 and 0.4 μs for DNA_{loop}, DNA_{mid} and DNA_{end} respectively. Combining the time-resolved results with IR absorption spectroscopy data indicates that the terminal position is likely to be substantially end-frayed in DNA_{end} and implies that the stem of the

DNA_{end} hairpin is essentially comprised of just three formal base pairs. The consequence of this is a less stable structure that inevitably dissociates on a shorter timescale. Conversely, the central AT pair in DNA_{mid}, which showed strong IR spectroscopy evidence for being in a W-C H-bonded state via prominent bands associated with base paired AT vibrational modes, melts on a longer timescale than that in DNA_{end}, as would be expected. There is however evidence from the shorter melting time for the sequence relative to DNA_{all} that the AT pair in this mid position lowers the stability of its neighbours. The nearest neighbour (NN) method predicts the difference in enthalpy to be as much as 24.7 $kJmol^{-1}$ in dsDNA, though this is not recovered in our hairpin samples from either the melting point or the activation energy for this sequence.[27–30]

The loop closing AT in DNA_{loop} does not show evidence from the IR absorption data for fraying like that found when the label is placed at the stem terminal end. The melting data support this through observation of the longest melting timescale for the A_R mode of DNA_{loop}. Although it will lower the stability of its neighbouring GC, this base pair will be influenced through base stacking by a sequence of three GC pairs in the remainder of the stem. Overall, the data combine to suggest that the DNA hairpins unfold from the frayed terminal end, though it is hard to rule out whether this is a consequence of the fact that the terminal AT base pair is more unstable, effectively creating an initiation point for unfolding. It is however noteworthy that the analogous terminal AU pair in RNA_{end} shows no evidence of fraying.

Turning to the refolding dynamics shows that the DNA hairpin which refolds with the shortest timescale is DNA_{end}, then DNA_{mid} and finally DNA_{loop}. In the case of DNA_{end}, the implication is that refolding is driven by the formation of the sequence of three GCs near the loop. At the other end of the stem, DNA_{loop} showed the longest refolding timescale by a significant degree, which would

tend to suggest that the base pair at the base of the loop is important for nucleation of the refolding process, consistent with the shorter folding times for both DNA_{mid} and DNA_{end}, which both feature GCs in the 'loop' position. Taken together, the data point towards DNA hairpins which unfold from the terminal end, but refold from the loop.

4.4.4 RNA-DNA comparisons

It is interesting to note the differing impacts of site-specific AU/AT base pairs on the stems of the RNA and DNA hairpin loops. The dynamic data is consistent with a model whereby the RNA stem melts from the loop, with the terminal AU pair showing a propensity to remain in a W-C paired state after the bulk of the stem has melted. When RNA refolds, it appears to do so from the terminal end, with the placement of an AU pair in the end position significantly impairing refolding. In contrast, a similar analysis leads to the reverse model for DNA in which melting begins from the frayed end base pair, while the base pair at the closing point of the loop is most influential in the refolding process, behaving as the anchor point for reversing change.

At room temperature, two sequences showed a departure from an archetypal stem-loop configuration, RNA_{loop} and DNA_{end}, with both showing IR absorption spectroscopy evidence for structural disruption. This was most marked for the DNA_{end} sequence. Since, as discussed above, no fraying is observed in RNA_{end} (in contrast to DNA_{end}) this shows that the RNA structure, through its closer base-stacking and ordered backbone hydration is better able to stabilise the terminal AU pair.[23,31,32]

The A-form helical backbone structure that RNA adopts, which provides additional stability through closer base stacking and additional hydration of the backbone sugars, is likely to be responsible for the overall slower melting

dynamics of RNA. [23,31,32] The data suggests that the breakdown of base stacking is an important step in the melting of the double stranded stem in both DNA and RNA.[24–26] It is noteworthy that for R/DNA$_{mid}$, the only sequence without three consecutive GC base pairs, melting is not significantly accelerated in RNA or DNA relative to R/DNA$_{end/loop}$, suggesting that the collective nature of the base paired stem overcomes any breaking of a sequence of GCs by a central AU/AT pair. The enthalpy penalty for replacing a GC with an AU in a GC sequence is ~33 kJ mol^{-1} in RNA, as determined by NN method, 5 kJ mol^{-1} more than the penalty for similarly substituting an AT in DNA, a difference that aligns with the closer (tighter) base stacking in RNA.[33]

From the perspective of refolding, placing the marker base pair in the centre of the stem leads to similar refolding dynamics for R/DNA$_{mid}$. This is consistent with previous studies of hairpins with all-GC stems and again shows that having a GC base pair either side of the marker position prevents any significant perturbation of the stem behaviour.[1]

It is interesting, given the consistency of the refolding dynamics of R/DNA$_{mid}$ that the respective activation energies for refolding were found to be -54.9 ± 4.3 and -41.6 ± 4.5 kJ mol^{-1}. The differences may reflect some impact of entropic contributions to refolding. The complexities of using activation energy parameters to predict refolding behaviour is also shown by RNA$_{loop}$ and DNA$_{loop}$. Whilst yielding similar activation energies (-39.0 ± 0.8 and -43.0 ± 1.9 kJ mol^{-1} respectively) these sequences showed dynamic differences, with DNA$_{loop}$ refolding on longer timescales than RNA$_{loop}$. This difference in timescale confirms that the rate of base-stacking in DNA$_{loop}$ is slowed appreciably and, based on structural differences, we tentatively ascribe the rate differences to the loop closing position being particularly important in the nucleation and zipping of the hairpin sequence in DNA. In RNA$_{loop}$, the IR data shows that the

closing pair is only weakly associated, and yet it still refolds much faster than the RNA_{end} sequence. We cannot eliminate the possibility that faster folding of RNA_{loop} may be due to it adopting a different motif as observed in previous work on UNCG loop closing pairs,[39] but for this particular sequence at least, the data implies that the loop closing position is less limiting in reforming the loop in RNA than DNA.

The relative behaviour of RNA_{end} and DNA_{end} is the most divergent, defining the extremes of the refolding lifetimes of all the sequences. The refolding lifetime of DNA_{end} is relatively short, being most similar to DNA_{all}, and this supports the view that it behaves chiefly as a stem of three GCs, since the AT is frayed at room temperature. This is clearly not the case for RNA_{end}, where the AU terminal bases appear relatively strongly paired and remain so for an appreciable period during stem melting.

The fact that replacing the RNA_{end} terminal base pair with an AU leads to slower refolding compared not just to DNA_{end}, but all the other sequences, indicates the different relative importance of the terminal position in refolding compared to DNA. Since the limiting step is believed to be formation of the first base-pair and subsequent stabilising base-stacking, this implies the RNA terminal position is important as a base-pair nucleation site in this sequence. In general, the probability of a stabilized base pair forming may be similar in RNA and DNA, but their different structures and flexibility may influence where that particular position is in the sequence; RNA appears to favour nucleation from the terminal end, while DNA prefers to refold from the loop closing position. The difference between their activation energies for refolding of -74.8 ± 0.4 and -24.0 ± 2.6 kJ mol^{-1} for RNA_{end} and DNA_{end} respectively is also indicative of the relative importance of these positions on the stability of their respective structures; they have the highest and lowest refolding activation energies of the sequences.

From a biological mechanism perspective, it is important to consider whether the nature of the loop structures may also play a role in determining the observed folding mechanism. Apart from closer stacking, the RNA A-form helix also has a wider helical diameter and adopts a slightly different loop motif. [23,34] The UNCG tetraloop in RNA forms a specialised "Z-turn", formed by a trans-sugar wobble base-pair interaction between the first and fourth bases, facilitated by the C2'-endo puckering of the third residue and the third and fourth ribose rings are configured in a head-to-tail orientation resulting in a O4'-π stacking contact (see Fig. 1.4 and 1.6 in Chapter 1).[35] This is a key stabilising structure, but it is relatively rigid and, as such, may favour nucleation of the stem at a point distant from the loop, as we observe.

In DNA by comparison, the third residue of the loop is able to freely form either the C2'-endo or C3'-endo conformations.[36] This may account for a greater flexibility in adjusting to a different geometry, allowing the stem to form from the base of the loop, but could also explain why the UNCG structure is less stabilising in DNA in general. It is noteworthy that in nature 60% of RNA tetraloops are closed by a CG pairs and 20% GC (80% by some combination of G and C).[37] The free energy difference ($\Delta\Delta G^{\circ}_{37}$) between the two is 7.8 kJ mol^{-1}, while a UA closing pair is 11.7 kJ mol^{-1}.[38] Some studies have determined that the difference in stability and geometry between even CG and GC closing pairs can cause the formation of different loop motifs altogether[39] supporting our observation that an AU pair imparts more flexibility into the stem structure in comparison to a GC or CG. Based on our observations, the dominant presence of GC or CG pairs at the closing point of naturally-occurring loops would not appear to be due to folding timescales, as the refolding dynamics of $RNA_{loop/mid/end}$ do not correlate with having a GC or AT in this key position.

4.5 Conclusions

We have demonstrated that using temperature-jump/drop IR spectroscopy alongside IR absorption measurements allows the impact of stem base sequence upon melting and refolding dynamics of oligonucleotide hairpins to be measured. The melting timescales of the RNA and DNA hairpins highlight key differences between them. The DNA hairpins are less stable than analogous RNA sequences, with insertion of an AT at the terminal end causing the DNA hairpin to dissociate on the shortest timescale, with the longest melting timescale observed when the marker bases were placed in the loop closing position. This is consistent with evidence for fraying of the terminal base pair and suggestive of a melting mechanism which progresses towards the loop. For RNA, melting timescales were found to be largely insensitive to stem base sequence, however, placement of an AU at the terminal position led to a long melting timescale, reminiscent of bubble formation, which we ascribe to increased conformational flexibility of the AU base pair and indicated a melting direction that proceeds away from the loop.

Refolding dynamics showed that DNA and the equivalent RNA sequence generally refold on similar timescales, however placement of the AU/AT in the RNA_{end} and DNA_{loop} positions significantly impaired the refolding pathway, leading to longer refolding timescales. These results suggest that the folding pathway in RNA initiates at the terminal end and propagates toward the loop, while in the case of DNA there is a preference toward zipping from the loop closing base pair. Taken together, the melting and refolding data for DNA and RNA sequences shows that the mechanisms differ significantly, and we attribute this to stronger base stacking in RNA which affects melting while greater flexibility of single-stranded DNA sequences gives access to different transient structures enroute to refolding.

4.6 Appendix

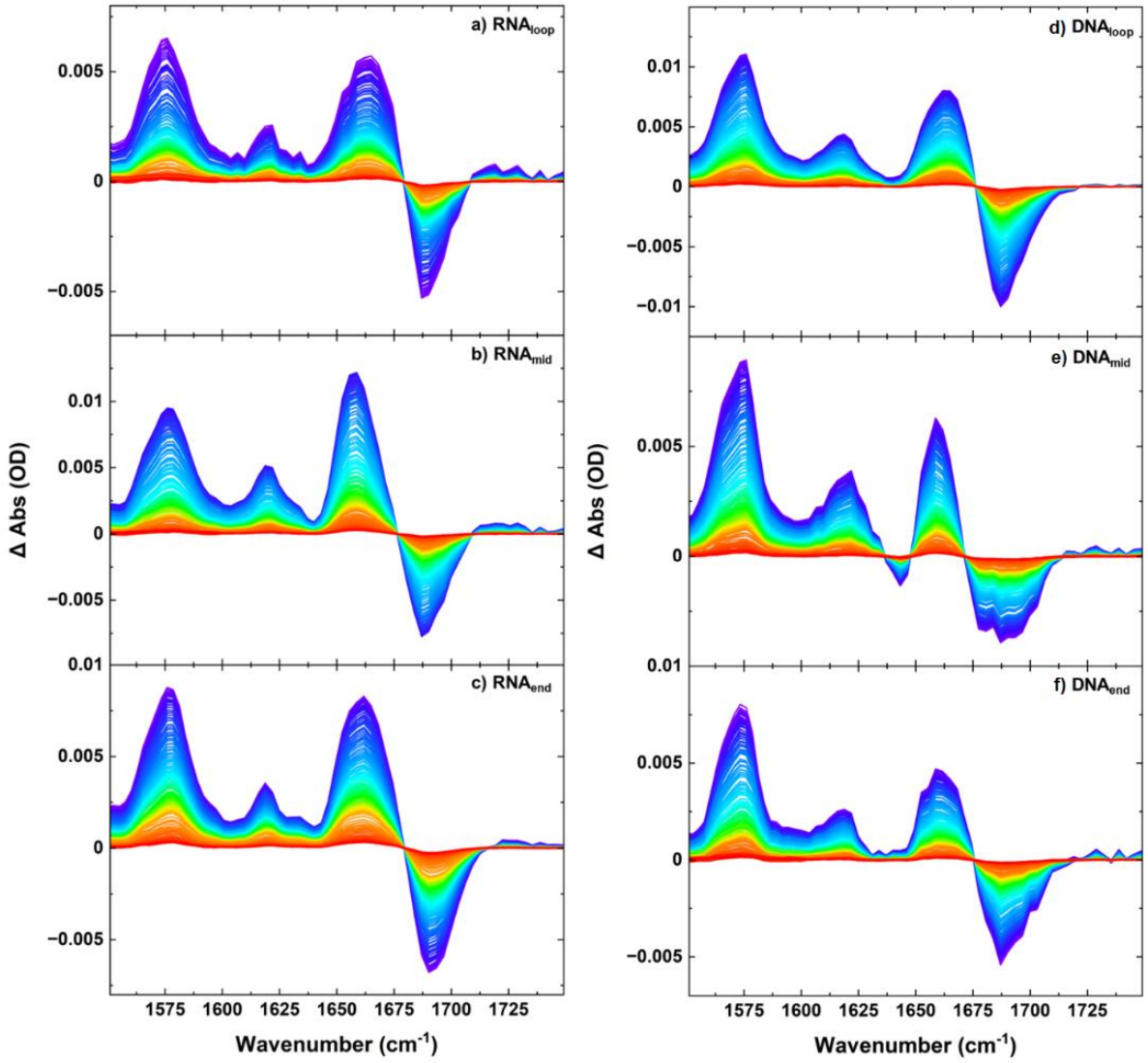

Figure A4.1: *T-jump decay spectra for a) RNA$_{loop}$, b) RNA$_{mid}$, c) RNA$_{end}$ and d) DNA$_{loop}$, e) DNA$_{mid}$, f) DNA$_{end}$ showing the decay of the hairpin signal from the peak signal (blue) to 4 ms (red). Spectra were taken at a T_0 nearest to T_m – 5 °C for each. T-jump spectra are shown as a pump on – pump off difference spectra with the increase in amplitude of a band represented as a positive peak. Spectra were baseline corrected for visual clarity.*

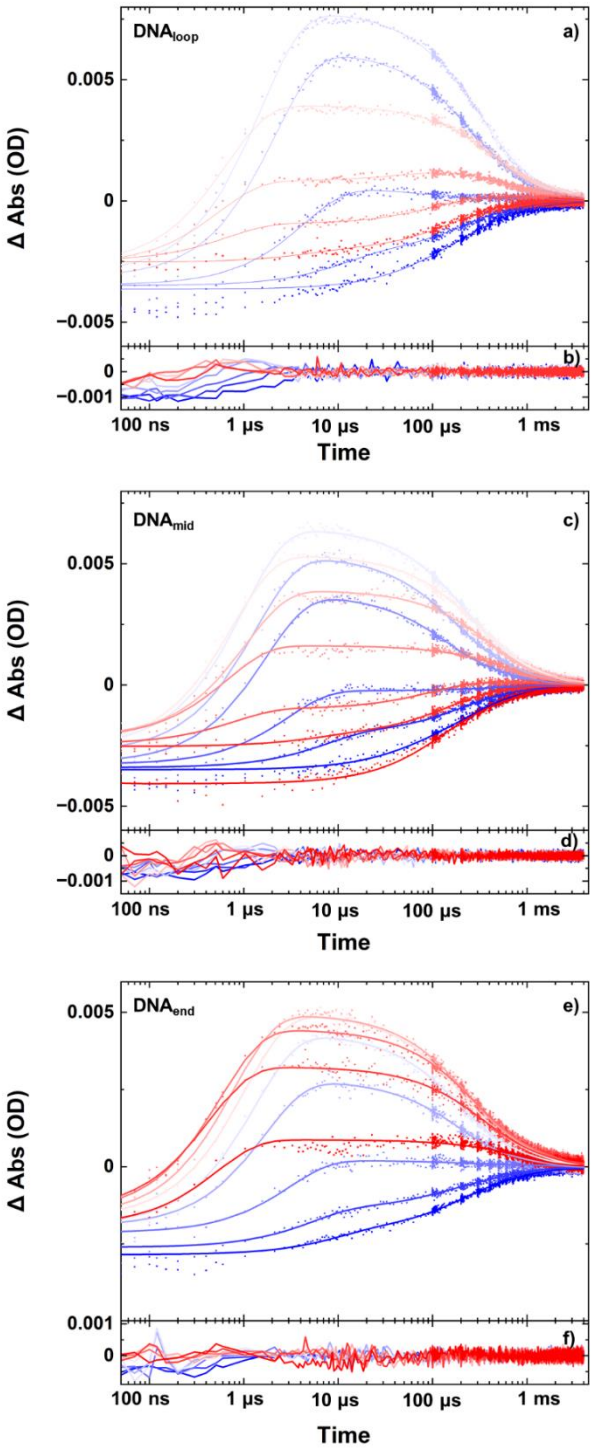

Figure A4.2: *T-jump kinetics showing Temperature and time-dependence of the G_R of the DNA sequences a) DNA_{loop}, c) DNA_{mid} and e) DNA_{end}. Data (dots) are shown from a T_0 of 20 °C to 80 °C (blue-red) along with the triple exponential fits (lines). Temperature dependent kinetic fit residuals are shown to scale below their respective kinetics (b, d, f).*

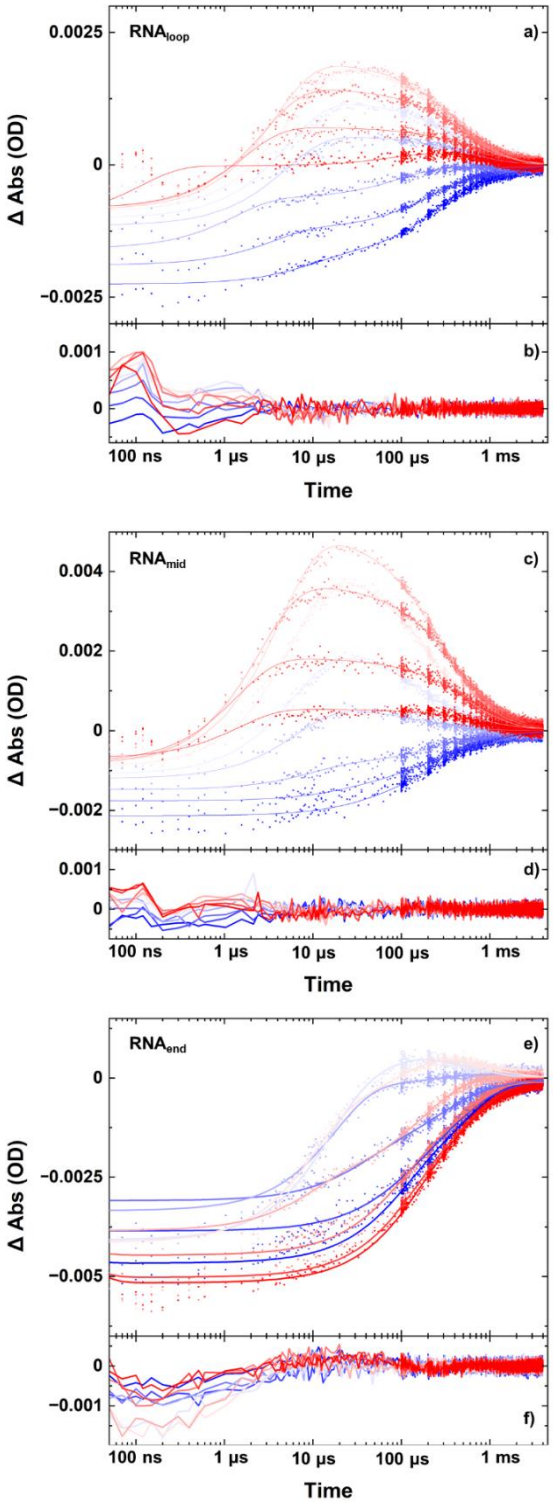

Figure A4.3: *T-jump kinetics showing Temperature and time-dependence of the A_R of the RNA sequences a) RNA$_{loop}$, c) RNA$_{mid}$ and e) RNA$_{end}$. Data (dots) are shown from a T_0 of 20 °C to 80 °C (blue-red) along with the triple exponential fits (lines). Temperature dependent kinetic fit residuals are shown to scale below their respective kinetics (b, d, f).*

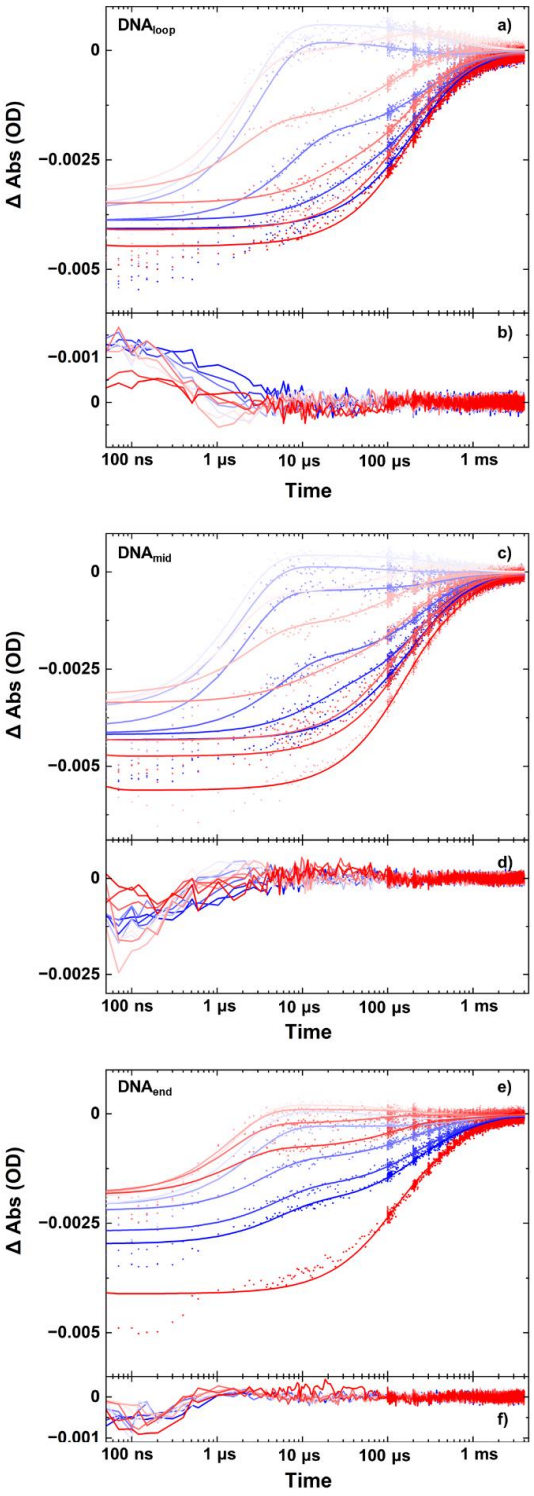

Figure A4.4*: T-jump kinetics showing Temperature and time-dependence of the A_R of the DNA sequences a) DNA_{loop}, c) DNA_{mid} and e) DNA_{end}. Data (dots) are shown from a T_0 of 20 °C to 80 °C (blue-red) along with the triple exponential fits (lines). Temperature dependent kinetic fit residuals are shown to scale below their respective kinetics (b, d, f).*

4.7 References

1 C. P. Howe, G. M. Greetham, B. Procacci, A. W. Parker and N. T. Hunt, *J. Phys. Chem. Lett.*, 2022, **13**, 9171–9176.

2 R. Fritzsch, G. M. Greetham, I. P. Clark, L. Minnes, M. Towrie, A. W. Parker and N. T. Hunt, *J. Phys. Chem. B*, 2019, **123**, 6188–6199.

3 G. M. Greetham, I. P. Clark, B. Young, R. Fritsch, L. Minnes, N. T. Hunt and M. Towrie, *Appl. Spectrosc.*, 2020, **74**, 720–727.

4 L. Minnes, G. M. Greetham, D. J. Shaw, I. P. Clark, R. Fritzsch, M. Towrie, A. W. Parker, A. J. Henry, R. J. Taylor and N. T. Hunt, *J. Phys. Chem. B*, 2019, **123**, 8733–8739.

5 M. Banyay, M. Sarkar and A. Gräslund, *Biophys. Chem.*, 2003, **104**, 477–488.

6 J. Dale, C. P. Howe, H. Toncrova, R. Fritzsch, G. M. Greetham, I. P. Clark, M. Towrie, A. W. Parker, T. C. McLeish and N. T. Hunt, *Phys. Chem. Chem. Phys.*, 2021, **23**, 15352–15363.

7 P. J. Sanstead and A. Tokmakoff, *J. Phys. Chem. B*, 2018, **122**, 3088–3100.

8 P. J. Sanstead, P. Stevenson, A. Tokmakoff and A. Tokmako, *J. Am. Chem. Soc.*, 2016, **138**, 11792–11801.

9 C. Lee, K. H. Park and M. Cho, *J. Chem. Phys.*, 2006, **125**, 114508.

10 C. Lee and M. Cho, *J. Chem. Phys.*, 2006, **125**, 114509.

11 C. Lee, K.-H. Park, J.-A. Kim, S. Hahn and M. Cho, *J. Chem. Phys.*, 2006, **125**, 114510.

12 C. Lee and M. Cho, *J. Chem. Phys.*, 2007, **126**, 145102.

13 A. L. Stancik and E. B. Brauns, *J. Phys. Chem. B*, 2013, **117**, 13556–13560.

14 M. Abdelkafi, N. Leulliot, V. Baumruk, L. Bednárová, P. Y. Turpin, A. Namane, C. Gouyette, T. Huynh-Dinh and M. Ghomi, *Biochemistry*, 1998, **37**, 7878–7884.

15 C. S. Peng, K. C. Jones and A. Tokmakoff, *J. Am. Chem. Soc.*, 2011, **133**, 15650–15660.

16 L. E. Revell and B. E. Williamson, *J. Chem. Educ.*, 2013, **90**, 1024–1027.

17 J. A. Wyer, M. B. Kristensen, N. C. Jones, S. V. Hoffmann and S. B. Nielsen, *Phys. Chem. Chem. Phys.*, 2014, **16**, 18827–18839.

18 T. E. Ouldridge, P. Šulc, F. Romano, J. P. K. Doye and A. A. Louis, *Nucleic Acids Res.*, 2013, **41**, 8886–8895.

19 M. C. Lin and R. B. Macgregor, *Biochemistry*, 1997, **36**, 6539–6544.

20 A. Ansari, S. V. Kuznetsov and Y. Shen, *Proc. Natl. Acad. Sci. U. S. A.*, 2001, **98**, 7771–7776.

21 M. I. Wallace, L. Ying, S. Balasubramanian and D. Klenerman, *Proc. Natl. Acad. Sci. U. S. A.*, 2001, **98**, 5584–5589.

22 R. J. Menssen and A. Tokmakoff, *J. Phys. Chem. B*, 2019, **123**, 756–767.

23 I. Anosova, E. A. Kowal, M. R. Dunn, J. C. Chaput, W. D. V. Horn and M. Egli, *Nucleic Acids Res.*, 2016, **44**, 1007–1021.

24 W. Zhang and S. J. Chen, *Biophys. J.*, 2006, **90**, 765–777.

25 K. Sarkar, K. Meister, A. Sethi and M. Gruebele, *Biophys. J.*, 2009, **97**, 1418–1427.

26 K. Sarkar, D. A. Nguyen and M. Gruebele, *RNA*, 2010, **16**, 2427–2434.

27 J. SantaLucia, H. T. Allawi and P. A. Seneviratne, *Biochemistry*, 1996, **35**, 3555–3562.

28 H. T. Allawi and J. SantaLucia, *Nucleic Acids Res.*, 1998, **26**, 2694–2701.

29 J. SantaLucia, *Proc. Natl. Acad. Sci. U. S. A.*, 1998, **95**, 1460–1465.

30 J. SantaLucia and D. Hicks, *Annu. Rev. Biophys. Biomol. Struct.*, 2004, **33**, 415–440.

31 E. M. Bruening, J. Schauss, T. Siebert, B. P. Fingerhut and T. Elsaesser, *J. Phys. Chem. Lett.*, 2018, **9**, 583–587.

32 A. Kundu, J. Schauss, B. P. Fingerhut and T. Elsaesser, *J. Phys. Chem. B*, 2020, **124**, 2132–2138.

33 T. Xia, J. SantaLucia, M. E. Burkard, R. Kierzek, S. J. Schroeder, X. Jiao, C. Cox and D. H. Turner, *Biochemistry*, 1998, **37**, 14719–14735.

34 G. P. H. Santini, C. Pakleza and J. A. H. Cognet, *Nucleic Acids Res.*, 2003, **31**, 1086.

35 L. Dascenzo, F. Leonarski, Q. Vicens and P. Auffinger, *RNA*, 2017, **23**, 259–269.

36 V. P. Antao, S. Y. Lai and I. Tinoco, *Nucleic Acids Res.*, 1991, **19**, 5901–5905.

37 C. R. Woese, S. Winker and R. R. Gutell, *Proc. Natl. Acad. Sci. U. S. A.*, 1990, **87**, 8467–8471.

38 J. M. Blose, D. J. Proctor, N. Veeraraghavan, V. K. Misra and P. C. Bevilacqua, *J. Am. Chem. Soc.*, 2009, **131**, 8474–8484.

39 M. Nakano, E. M. Moody, J. Liang and P. C. Bevilacqua, *Biochemistry*, 2002, **41**, 14281–14292.

5. Temperature-Jump 2D-IR of an RNA UNCG Tetraloop Hairpin

5.1 Introduction

To build on the work outlined in Chapter 3 and 4, the work conducted in this chapter uses non-equilibrium T-jump experiments with 2D-IR probing to further investigate the interactions and dynamics that occur in RNA hairpin loops. Equilibrium 2D-IR was also used to provide a benchmark comparison and allow clear cross-peak structure assignments in order to identify coupling signatures, with changes noted between the folded hairpin and melted forms. To evaluate the advantages of 2D-IR compared to 1D methods, this chapter revisits the UNCG RNA_{all} hairpin sequence with a four GC stem and the analogous DNA_{all} hairpin.

Two-dimensional Infrared Spectroscopy is a non-linear optical spectroscopy that allows the measurement of ultrafast (sub-picosecond) vibrational dynamics with frequency dispersed pump and probe enabling the determination of coupling and energy transfer between vibrational modes.[1] By spreading the spectral information across a second frequency access, the vibrational dynamics of the different modes is correlated to each other, allowing the structural relationships in the nucleic acids to be determined. It also provides additional fine detail compared to 1D spectroscopy methods, due to supressing the underlying solvent signal and producing spectral bands with narrower linewidths, making it easier to assign complex spectral features. 2D-IR has been proven to be an effective method for the investigation of DNA vibrational dynamics,[2–14] elucidating dynamics such as the rapid energy transfer between the bases and the phosphate backbone enabling dissipation of excess energy in both double stranded and single stranded DNA.[9] By contrast very few 2D-IR studies have

focused on RNA. This is slowly changing, with recent work by Elsaesser, Fingerhut and co-workers focusing on hydration dynamics of the RNA and DNA backbone.[15] This demonstrated ordered hydration of the RNA backbone leading to additional coupling features in RNA, which influences the conformation and stability of the structure. The impact of hydration on melting was also investigated which showed that this ordered hydration breaks down on melting, resulting in a "DNA-like phosphate group hydration".[16] This was expanded on with several studies on the influence of ion interactions on the backbone dynamics.[17–20] The RNA base uracil has also been characterised with 2D-IR in a study on the nucleic acid bases.[6] While progress is being made, RNA adopts varied functional structures that have yet to be characterised by 2D-IR, including the vibrational dynamics of tetraloop hairpins.

Non-equilibrium 2D-IR measurements were first reported in 2004 to study metal carbonyl complexes using UV excitation to probe electronic excited states and photochemical processes.[21,22] Subsequently its use was extended to a range of inorganic systems.[23,24] This made it possible to gain the benefits of 2D-IR, correlation of sub-picosecond time resolution of vibrational mode dynamics, during a transitional process.

More recently, the dynamics of bio-molecular systems have been investigated using T-jump 2D-IR.[7,22,25–27] This has been led by Tokmakoff and co-workers who have applied T-jump 2D-IR to the study of a range of biological phenomena, including the dissociation dynamics of insulin dimers,[27] and the folding and unfolding of an intercalated DNA i-motif and ubiquitin protein.[7] This included the introduction of a continuous wave laser after the initial T-jump pulse in order to maintain the temperature to allow the system to fully respond to the change in the environment, before turning it off and allowing the system to return to its

initial equilibrium state. This allowed the measurement of dynamic processes over 100s of milliseconds allowing longer dynamics to be observed as well as controlling when cooling starts. The Tokmakoff group have also employed a T-jump dispersed vibrational echo method to study the influence of sequence, pH and methylation on DNA melting and folding dynamics.[28–30]

The work presented in this chapter uses T-jump 2D-IR to deepen our understanding of the conformational dynamics of RNA UNCG tetraloops through the melting process.[31] The T-jump studies in Chapter 3 and 4 on RNA and DNA hairpin melting dynamics determined that the RNA hairpins melted an order of magnitude slower compared to the equivalent DNA. This large discrepancy in behaviour was caused by the propensity for RNA to form A-form helices, allowing closer stabilising π-stacking of the base rings and greater hydration of the backbone.[15,32–34] Additionally, by including an AU/T marker at various points in the stem, it was determined that the melting of the DNA and RNA hairpins likely proceeds by different pathways, with the RNA hairpins melting from the loop closing base pair first and unzipping down the stem, proceeding via a "bubble" formation before full dissociation (see Chapter 4). The DNA hairpin by contrast tended to unzip from the terminal base pair of the stem toward the loop. By extending work on these systems into 2D-IR and T-jump 2D-IR, greater insight can be gained on the structural dynamics underpinning the functions of RNA and its different behaviour compared to DNA.

5.2.1 Two-Dimensional Infrared Spectroscopy

Two dimensional Infrared spectroscopy, similar to two dimensional NMR, spreads the linear IR spectrum over two frequency axes by exciting each mode and thus correlating the relationships between them, making it possible to observe coupling between modes, energy transfer and the influence of intermolecular interactions. Coupling and energy transfer pathways can be related to changes in the structure that are reflected in their spectral signatures and are challenging to resolve in 1D spectra as the features overlap.

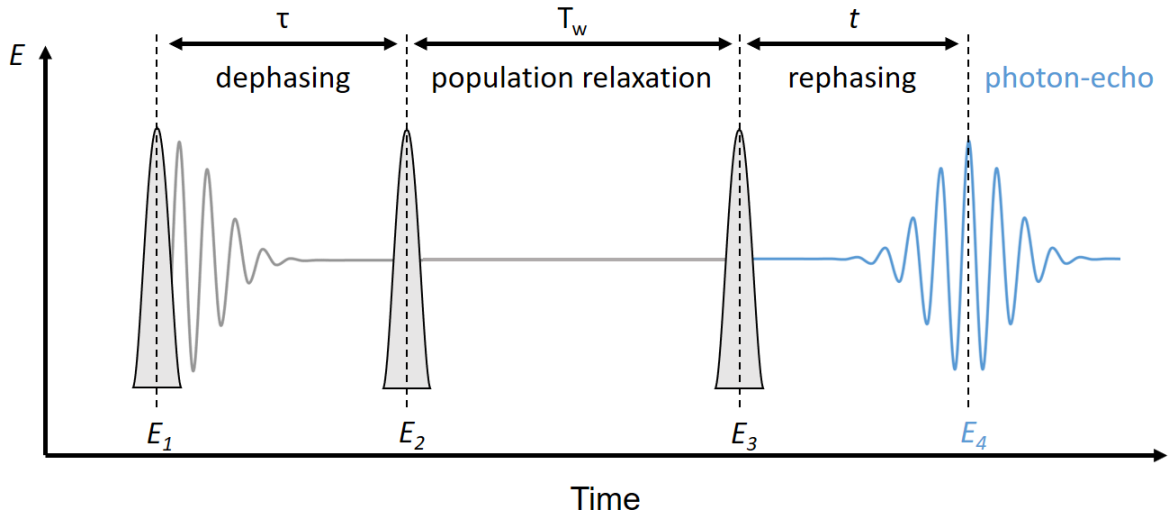

Figure 5.1: *Pulse sequence in a time-domain 2D-IR experiment, with two pump pulses (E_1 and E_2) and a probe pulse (E_3) causing the rephasing and emission of a photon-echo (E_4).*

Practically, 2D-IR spectroscopy is a "four-wave mixing" technique in which a signal is generated by the interaction of three laser pulses, E_1, E_2, and E_3 (Fig. 5.1) with the sample.[26] The emitted signal field, or "photon echo", is E_4. The time between E_1 and E_2 is the first coherence time, τ. The second time interval between E_2 and E_3 is T_w, the waiting time and the time between E_3 and E_4 is the second coherence time, *t*.

The first pulse, E_1, resonant with the v=0-1 transition creates a coherent superposition of the ground and first excited state (Fig. 5.2), generating a macroscopic polarization. The superposition oscillates at the frequency of the v = 0-1 transition, defining the pump frequency coordinate on the 2D-IR plot. Due to the high field intensities of the laser pulses a third order polarisation is generated:

$$\mathrm{P(t)} = \epsilon_0 \left(\chi^{(1)}E(t) + \chi^{(2)}E^2(t) + \chi^{(3)}E^3(t)\right) \qquad 5.1$$

The macroscopic polarization is determined by the permittivity of free space, ϵ_0, the electric susceptibility, χ, and the oscillating electric field, E(t). The terms $^{(2)}$, $^{(3)}$ are the non-linear susceptibilities and have decreasing magnitude as their order increases. The second order term is generally not considered as it is zero for isotropic media.[35]

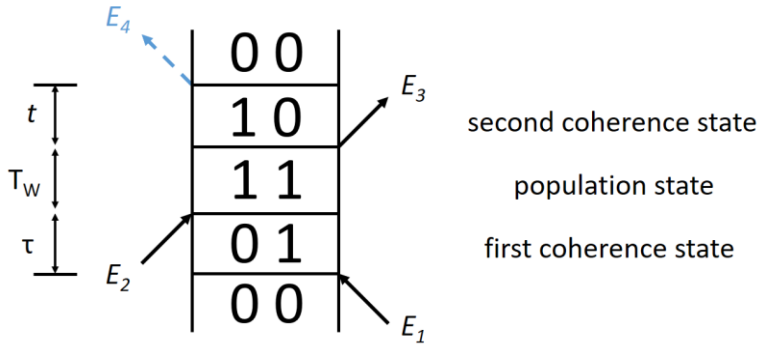

Figure 5.2: *Diagram showing one of many possible pathways for the sequence of pump and probe pulses in a 2D-IR experiment, each corresponding to a peak in the 2D-IR spectrum.*

The coherence caused by the first pulse dephases over time. The second pulse, E_2, creates a population state, either of the ground or first excites state (Fig. 5.2). During the waiting time, T_w, after the second pulse, is the interval in which vibrational dynamics are observed. By taking successive measurements with varying T_W the spectral diffusion (discussed further below) and vibrational relaxation can be measured.

After the waiting time E_3 returns the system to a superposition of states which will depend on the particular population state after the first coherence time. The oscillation frequency of this superposition determines the probe frequency of a feature in the 2D-IR spectrum. The oscillators are not initially in phase, but the third pulse causes a rephasing process that returns the system to coherence after a time such that $t = \tau$. With the return of the coherence state, the corresponding macroscopic polarization is recovered and the photon echo emitted. For rephasing pathways inhomogeneous broadening is also reversed, making it possible to observe the homogeneous broadening in isolation. This is possible because resonant frequencies responsible for the inhomogeneous rephasing remain constant. While they oscillate with varied frequencies, covering different distances over the dephasing time, when the third pulse reverses them, they travel the same distance back which results in the re-emergence of the macroscopic polarization after the same interval of time, namely τ.[35]

Structure of 2D-IR Spectra

The features of a 2D-IR spectrum can be best illustrated by use of a model molecular system, a pair of oscillators coupled via mechanical or electrostatic interactions, and by relating those features back to an energy level diagram of the transitions they pertain to. The IR absorption spectrum for two coupled vibrational modes has two clearly distinct peaks (Fig. 5.3, a), but coupling information cannot be easily extracted.[26]

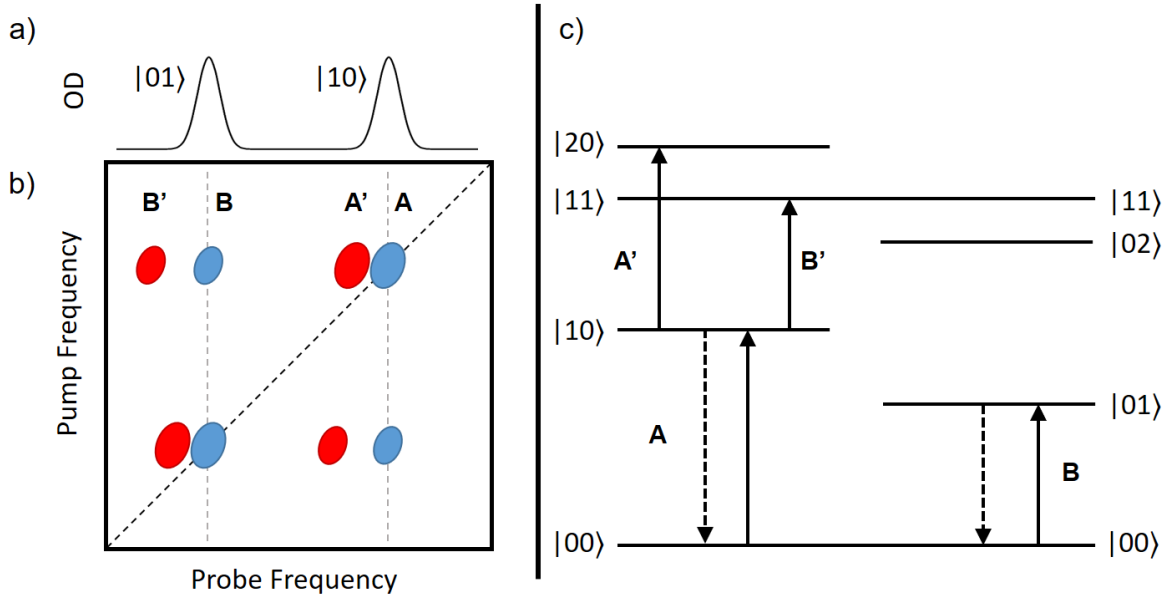

Figure 5.3*: Features of a 2D-IR spectrum, showing a) the FT-IR spectrum of two coupled modes with bleaches shown in blue and absorption peaks in red, b) 2D-IR spectrum of two coupled modes, and c) corresponding energy level diagram for the features of the 2D-IR.*

The 2D-IR (Fig. 5.3 (b)) contains much more information as a direct consequence of the three pulses allowing the excitation of multiple vibrational transitions and resulting in a complex sequence of interactions, populating vibrational levels up to v=2. The 1D features appear on the diagonal while the off-diagonal features are representative of coupling and energy transfer processes. The peaks seen in the 2D-IR can be assigned according to the energy level diagram shown in Figure

5.3 (c). For the peak labelled A, the first pulse E_1 causes a coherence state, in this case of the ground state and the first exited state. Using notation for two modes, this is a coherence between $|00\rangle$ and $|10\rangle$ of the coupled modes (where $|01\rangle$ corresponds to the first excited state of the second oscillator). The incidence of E_2 then results in a population state of $|10\rangle$ and E_3 creates a second coherence state between $|00\rangle$ and $|10\rangle$. This results in the photon echo having both pump and probe frequencies corresponding to the $|00\rangle$ - $|10\rangle$ transition, with a negative peak due to the phase relationship between the first and second coherence states. This is the pathway seen in Fig. 5.2. In the photon echo description, the features of 2D-IR are a consequences of different pathways. However, the pump probe description makes the relationships between them more apparent and will be used going forward.

Using pump-probe terminology to describe the same process that results in A, an initial pump resonant with the A transition, excites some population from the $v = |00\rangle$ to $v = |10\rangle$ state. The probe pulse then records a reduced absorption of the $|00\rangle$ - $|10\rangle$ transition, leading to negative peak, a bleach, at the frequency of the $|00\rangle$ - $|10\rangle$ transition. The probe can also result in stimulated emission of the the$|10\rangle$ state contributing to this feature, or further excite the population of $v = |10\rangle$ to $v = |20\rangle$, corresponding to the transition A'. A' is displaced to lower frequency compared to A by the anharmonicity of the transition. Without anharmonicity, the transitions between A and A' would be of equal magnitude and frequency and would cancel each other out.

In addition to the peak pairs on the diagonal, there are also peaks that appear in the off-diagonal regions, called cross-peaks, here labelled B and B'. The cross-peak labelled B is a negative feature, which is caused by coupling between the two modes and is a consequence of their shared ground state. In other words,

when A is bleached, so is B. Exciting one also influences the other, changing its frequency, resulting in a combination band mode, the |11⟩ state. This is another possible transition from the first excited state |01⟩ after E_2, so it appears as a positive peak, the cross peak B' at a pump frequency corresponding to the |00⟩ - |10⟩ transition and probe frequency corresponding to |01⟩ - |11⟩. It is also shifted compared to B by the anharmonicity of the |01⟩ - |11⟩ transition, called the off-diagonal anharmonicity. Off-diagonal peak intensities are also important as they can indicate the distance and angle between the coupled dipoles. The lower peaks on the 2D-IR are the direct corollary, in the case where the pump frequency is tuned to the resonant frequency of the second of the two modes.

Spectral Diffusion

For a single oscillator the linewidth of its peak would be Lorentzian and purely determined by the homogeneous broadening of the vibrational mode. However in solution an ensemble of oscillators result in inhomogeneous line broadening along the diagonal of the 2D-IR, producing an elliptical line shape (Fig. 5.4, (a)).

This lineshape evolves, becoming less elliptical as the frequencies of the oscillators changes over time in response to changes in the environment such as the sub-picosecond variation in the hydrogen bonding network in solution (Fig. 5.4, (b)). So by taking successive 2D spectra at different delays between the E_2 and E_3 pulses (the waiting time, T_W) this evolution in lineshape, known as the spectral diffusion can be measured. This is done by analysis of the change in the central line slope[36] (long axis of the ellipse) or nodal line slope[37,38] (contour between the v=0-1 and v=1-2 peaks, shown in Fig. 5.4)and the rate of change can be used to infer the environment of the oscillators.

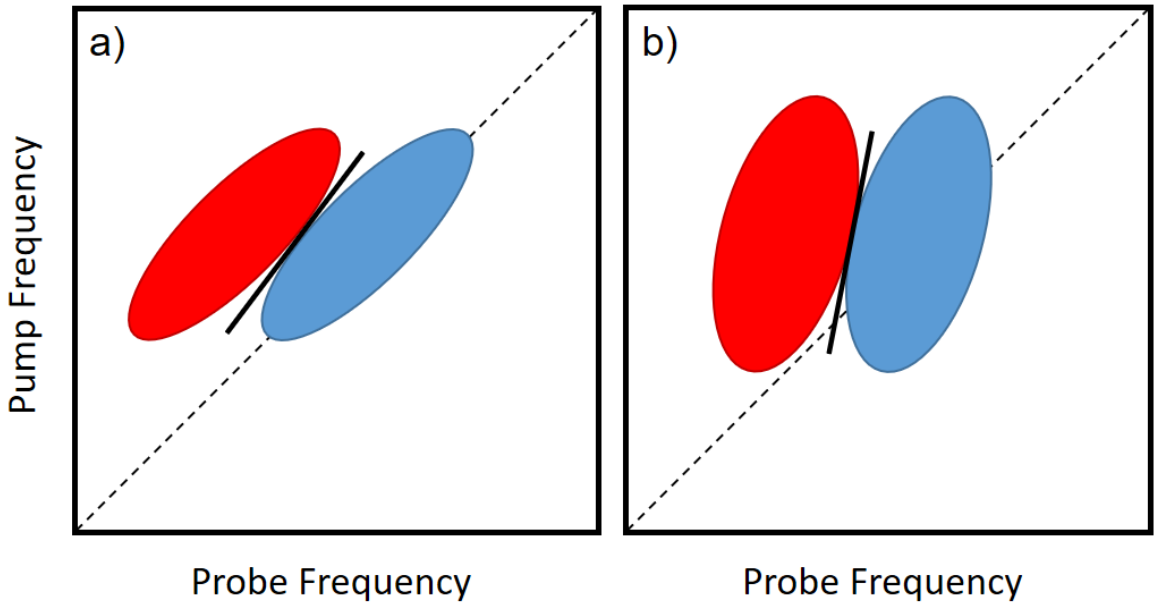

Figure 5.4: *Example 2D-IR spectra of a single vibrational mode recorded at a) T_W=0 (left) and b) T_W>0 (right). Increasing T_W results in broadening of the spectral features in the 2D-IR spectrum. The nodal line slope (black line) changes due to spectral diffusion.*

DNA and other nucleotide sequences have significantly more vibrational modes, which often overlap in the IR spectral region of interest, namely carbonyl, C=C and C=N IR modes.[39] The ability of 2D-IR techniques to resolve the details of such crowded spectral regions has allowed the complex dynamics of DNA to be resolved.[8,9,12,13] This has allowed coupling patterns and spectral diffusion to be used as a means to infer spatial and conformational relationships in the nucleic acid system. By Applying 2D-IR, understanding of the dynamics and interactions of RNA can be furthered in a similar way.

5.2.2 Temperature-jump 2D-IR

Chapters 3 and 4 implemented a temperature-jump method effective for directly measuring nucleic acid melting and refolding dynamics over the timescale of nanoseconds to milliseconds. Experiments presented in this chapter build on that work, using T-jump initiation followed by 2D-IR probing, which allows discrete 2D-IR spectra to be recorded at defined points through the dynamic process of nucleic acid strand melting.[7,22,25–27] This makes it possible to gain all the benefits of 2D-IR, determination of the structural arrangements and environmental, and observe their dynamic changes in real time through the melting process.

5.3 Experimental

The RNA and DNA oligomer sequences 5′-GCGC(XACG)GCGC-3′ (RNA: X = U; DNA: X = T; "(...)" indicates the loop position) were prepared as outlined in 2.1. These are designated RNA_{all} and DNA_{all}, consistent with previous work (see Chapter 3 and 4).

FT-IR Absorption measurements of the sequences were conducted over the temperature range from 20 – 80 °C through the melting transition with the spectroscopy and mode assignments discussed in detail in Chapter 3, §3.3.1.

The Temperature jump 2D-IR method was adapted using the STFC Central Laser Facility's ULTRA spectrometer T-jump method.[31,40–42] A detailed description of the T-jump set up is described in Ch. 2, §2.3.

Equilibrium 2D-IR measurements of RNA_{all} and DNA_{all} were taken at 20 and 80 °C and over a range of T_w from 250 fs to 3 ps. T-jump 2D-IR were taken at a T_0 of T_m – 5 °C to observe the maximum transition through the hairpin melting. For

RNA_{all} this was 75 °C and 70 °C for DNA_{all}. T-jump 2D-IR measurements were held at a 2D-IR T_W of 250 fs, while the spectra were recorded at T-jump delays of -20 ns, 500 ns and the peak melting time (6 µs for DNA_{all} and 30 µs for RNA_{all}).

5.3.1 Two-dimensional Infrared Spectrometer

2D-IR and T-jump 2D-IR experiments on RNA and DNA hairpins were performed using the STFC Central Laser Facility's ULTRA B spectrometer system.[43,44] The 2D-IR pulses are generated using the same method as described in Ch. 2, §2.2.

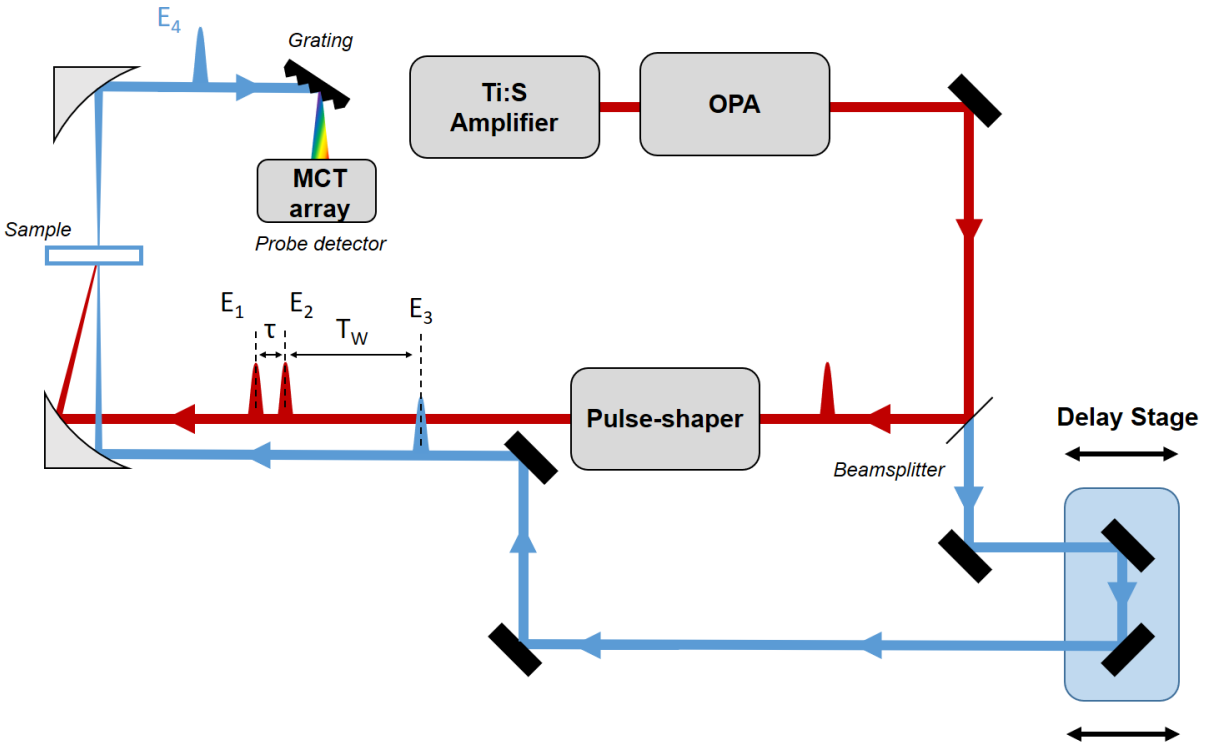

Figure 5.5: *ULTRA B 2D-IR experimental set up with a pulse-shaper for collection of time-domain 2D-IR spectra. The pump pulses, E_1 and E_2 are shown in red, separated by τ, controlled by the pulse-shaper. The probe pulse E_3 is shown in blue is separated from E_2 by T_w, determined by the delay stage. The emitted signal E_4 is emitted collinearly with the probe beam due to the pump-probe beam geometry.*

For the work conducted here, an OPA was used to produce pulses with a pulse duration of <50 fs and a repetition rate of 10 kHz and a bandwidth of 300 cm^{-1} centred at 1650 cm^{-1}, resonant with the nucleotide base vibrational modes. The 2D-IR experimental configuration is shown in Figure 5.5.

With a pulse generated with the desired Mid-IR centre frequency, the beam is put through a 90:10 splitter to create the pump and probe beams and the pump pulse is directed into the pulse-shaper (PhaseTech). As shown in Fig. 5.6, this pulse is dispersed on a grating before being collimated and passed through a germanium acousto-optic modulator (AOM). An acoustic wave, arbitrarily programmable with a computer controlled waveform generator, is passed through the AOM crystal, changing the refractive properties across the crystal perpendicular to the beam path, which acts as a selective grating to deflect frequencies with chosen amplitude and phase (5.6 (b)).[35,45] By precise manipulation of the acoustic waves, an interference pattern can be generated equivalent to a pulse-pair in the time domain (5.6 (c)).[35] The two collinear pump pulses are separated by a pump-pump delay, τ. By adjusting the interference pattern the τ delay can be scanned to generate the 2D spectrum.

The pump and probe pulse are focused on the sample with the transmitted signal light, collinear with the probe due to the pump-probe geometry, dispersed by a grating onto a 128-channel MCT detector. The intensity at each pixel is recorded as a function of τ and Fourier transformed to obtain the pump axis of the 2D-IR spectrum.

.

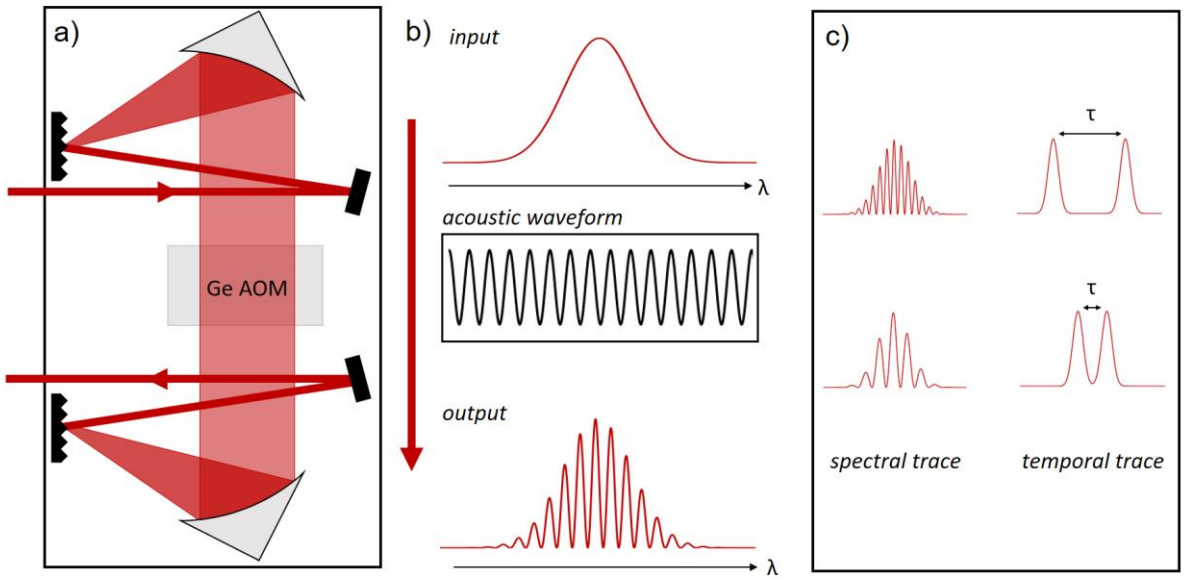

Figure 5.6: *Pulse-shaping using a Germanium Acousto-Optic Modulator (AOM). a) A schematic of the pulse-shaper, b) representation of arbitrary wave patterns passed through the AOM perpendicular to the beam which selectively refracts specific wavelengths to superimpose the wave on the beam frequency profile, producing a pair of pulses in the time-domain, and c) different spectral profiles corresponding to different time interval between pulses in the time-domain.*

The timing with respect to the delay between the second pump pulse and the probe pulse, the waiting time T_w is also strictly controlled using an optical delay stage. This is fixed for each 2D-IR spectrum, but scanning T_w over multiple 2D-IR spectra provides snapshots into the response of the excited state dynamics of the system over a defined time interval, making it possible to determine vibrational lifetimes and spectral diffusion

The generation of the collinear pump pulses using pulse-shaping has a number of benefits over the use of interferometric methods. Pulse-shaping improves data quality by generating both pulses with a single input, side stepping the need for slow delay stages, guarantees collinearity of the pump pulses, and allows

precise control of the relative phases of the two pump pulses, making phase cycling possible.

Phase cycling consists of alternating the offset of the pump pulse phases (Δϕ) between 0 and π (Fig. 5.7 (a)), half the period of the pump electric field oscillation. When the offset is π the sign 2D-IR signal becomes negative (Fig. 5.7 (b)), but this does not have any influence on other contributions to the signal, such as errant light from the laser pulses scattered by the sample in the direction of the emitted field adding random noise into the data, or systematic sources of noise.

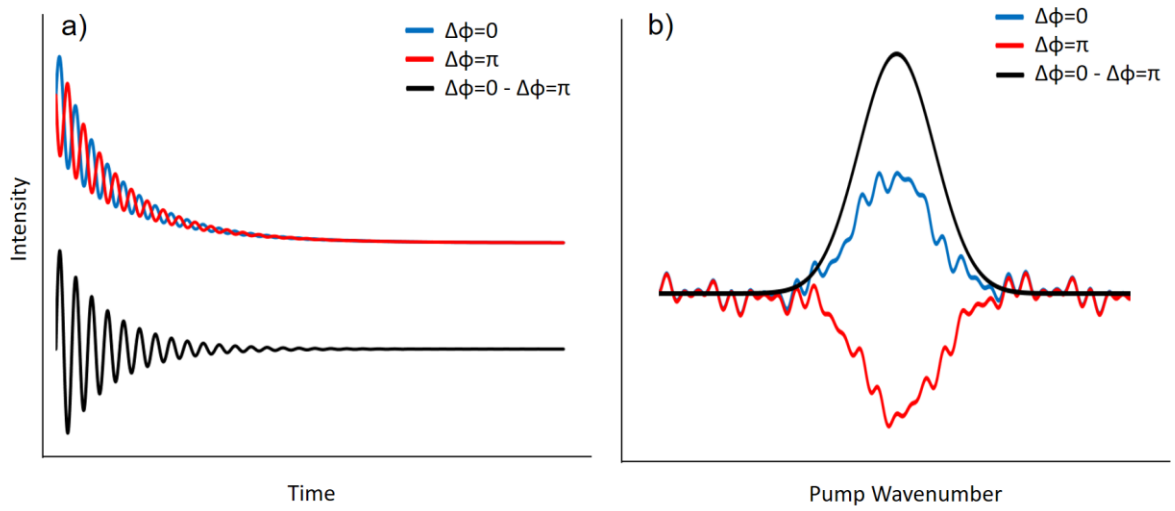

Figure 5.7: *Removal of background from a transient absorption using phase cycling a) Time scans of two signals with a phase difference between the two pump pulses of 0 and π (red and blue), and the subtracted product of the two (black). b) Representation of the Fourier transform of the time scans showing the signal enhancement and noise reduction as a result of subtraction of the two signals.*

Therefore, subtracting the two out of phase signals results in any other components of the light that are constant or periodic being suppressed, while the 2D-IR signal is increased, dramatically improving the signal to noise. This

allows for more clearly defined peak lineshapes and intensities, making it possible to determined spectral diffusion and anharmonicities much more easily and accurately.

T-jump 2D-IR Spectroscopy

The Temperature jump 2D-IR method was adapted from the T-jump and 2D-IR methods, with the T-jump pump generated as detailed in Ch. 2, §2.3.1 and 2D-IR pumps and probe pulses as described above.[31,40–42] To produce T-jump-2D-IR spectra, the delay time between the 2D-IR pump pulses (τ) was scanned while the T-jump-2D-IR delay time and the delay time between the 2D-IR pump and 2D-IR probe pulses (T_w) both remained fixed. By changing the T-jump pump-2D-IR probe delay time for each 2D-IR acquisition, T-jump-2D-IR spectra could be measured at time delays ranging from 1 ns to 4 ms following the T-jump.

5.4 Results

To gain deeper insight into the stability and melting mechanism of RNA UNCG tetraloops, the four GC stem UNCG tetraloop hairpin RNA_{all} and its TNCG DNA analogue DNA_{all} were studied using static temperature equilibrium 2D-IR in both single stranded, "melted" form and when annealed as a hairpin loop, as well as non-equilibrium T-jump 2D-IR through the melting transition to probe the dynamics of the melting process and track the coupling over time

5.4.1 Two-dimensional Infrared Spectroscopy

2D-IR spectroscopy was employed to capture the equilibrium spectral feature of RNA_{all} and DNA_{all}. These spectra were used to support the classification of the

base vibrational modes in conjunction with linear IR absorption spectra from 20 – 80 °C through the melting transition as studied previously and presented in Chapter 3, §3.3.1.[31] Compared to linear IR absorption methods, the third order non-linear response of 2D-IR enhances the signal from vibrational modes with relatively large extinction coefficients.[35,46–49] This results in the broadband solvent response in the 1650 cm^{-1} region to be suppressed and means solvent correction does not need to be applied. The primary focus of these measurements was the base region between 1550 and 1700 cm^{-1}. The frequency resolution of the peaks is improved in 2D-IR as well, due to narrower linewidths of the modes as a consequence of the non-linearity of the extinction coefficient dependence, allowing for greater distinction to be drawn between them in this convoluted base region.[35] For analytical purposes, the spectra taken with wait times of 250 fs will be used and compared, where the signals are strong and coupling and peak shapes more readily seen and interpreted.

5.4.2 RNA_{all} 2D-IR

Comparing the IR absorption spectrum of RNA_{all} (Fig. 5.8 (a, b)) to the on diagonal features apparent in the 2D-IR (Fig. 5.8, (c, d), black diamonds), a number of key modes can be identified. In the 20 °C 2D-IR spectrum (Fig. 5.8 (c)), the negative peak seen at 1580 cm^{-1} corresponds to the v=0-1 ground state bleach of the guanine ring vibrational modes, the G_R. The corresponding v=1-2 transient absorption peak of the G_R is also observed, shifted to lower frequency by the anharmonicity of the transition. A second pair of v=0-1 and v=1-2 peaks are seen around 1622 cm^{-1}, corresponding to the A_R.

The spectral region from 1630 to 1700 cm^{-1} has a greater number of contributions from different modes. A strong bleach is observed at 1686 cm^{-1} of

the G_SC_S(-), while the bleach seen at 1646 cm^{-1} corresponds to the G_SC_S(+), both with corresponding transient peaks at lower frequency.

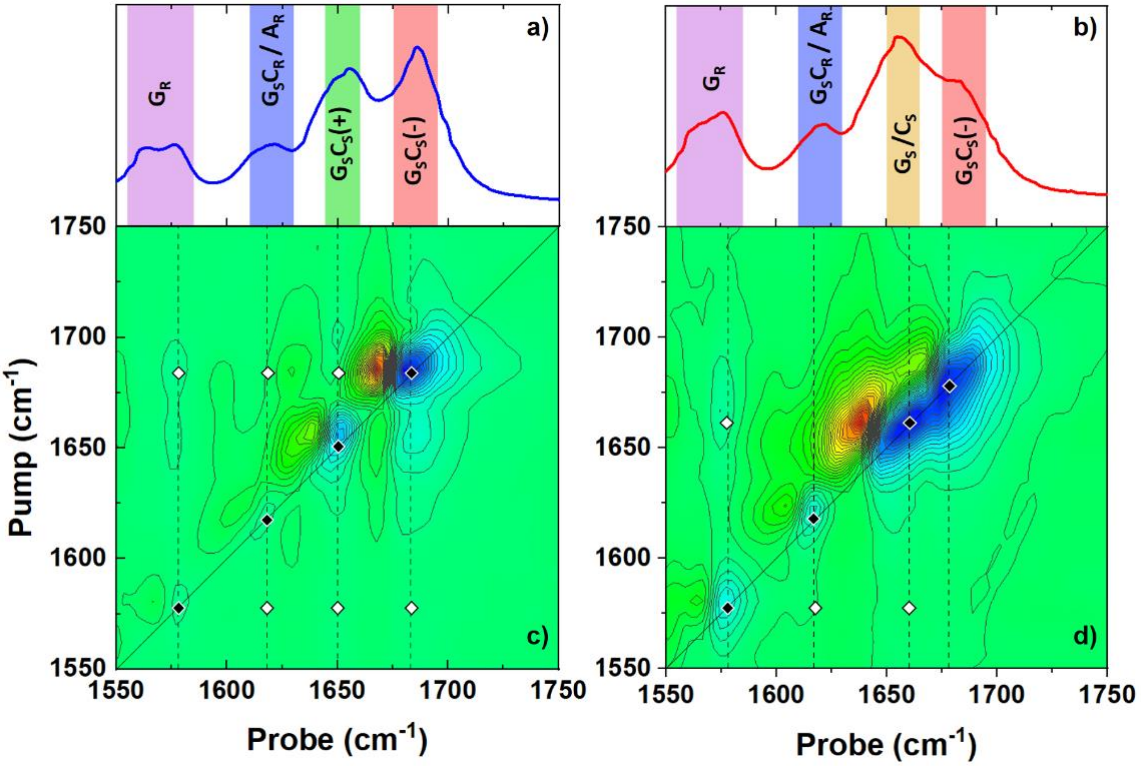

Figure 5.8: *IR spectra showing, FT-IR absorption spectra of RNA_{all} a) at 20 °C and b) 80 °C. Vibrational mode assignments are indicated by coloured panels:, G_R (1575 cm^{-1}, purple); overlapping A_R and G_SC_R (1620 cm^{-1}, blue); G_SC_S(+) (1648 cm^{-1}, green); overlapping G_S, C_S (1656 cm^{-1}, yellow); G_SC_S(-) (1686 cm^{-1}, red).*[39,50,51] *The notation used to identify the vibrational modes of the bases is consistent with that used in Refs 39, 50 & 51. Spectra have been solvent corrected. c) the 20 °C, 250 fs 2D-IR for RNA_{all}. d) the 80 °C, 250 fs 2D-IR for RNA_{all}. Overlaying guidelines indicate the position of modes identified in the IR absorption spectra. On diagonal features are indicated by black diamonds, while cross-peaks are shown by white diamonds.*

In the 80 °C 2D-IR spectrum (Fig. 5.8 (d)), the G_R mode is observed to increase in intensity through the melting transition, as discussed in Chapter 1, §1.2.1.[14,31,39,40,50–55] The A_R mode is not seen to change. Since there is only a single adenine present, unpaired in the loop and is not involved in the primary structural stabilisation of the UNCG Z-turn motif, it stays constant through the melting process.[56]

In general, the appearance of a broad bleach is apparent at 80 °C, a combination of the strong unpaired G_S mode at 1668 cm^{-1} and the C_S around 1650 cm^{-1} which shifts to higher wavenumber with temperature. Both are indicative of melting. Some degree of G_SC_S(+) and G_SC_S(-) is still visible as the melting temperature of RNA is quite high (75 °C). The melting curves and behaviour for RNA_{all} are discussed in greater detail in Chapter 3, §3.3.1.

The on diagonal features in the 2D-IR contour plot are apparent (Fig. 5.8 (b, d) black diamonds), and some coupling is clearly visible, but to properly analyse the more subtle coupling features it is often clearer to refer to the individual pump slices (Fig. 5.9), horizontal cross sections of the 2D-IR at a single pump frequency. All identified cross-peaks have been indicated in Figure 5.8 using white diamonds.

The Pump slices take at 1580 cm^{-1} for RNA_{all} (Fig 5.9 (a)), shows a clear G_R mode (purple panel) bleach of the v=0-1 transition and excitation of v=1-2, with an increase in signal observed when the sample is heated from 20 (black trace) to 80 ^oC (red trace), indicating a loss of base-stacking. Over the temperature change there is also a noticeable change in the coupling pattern. At 20 ^oC, there is clear coupling to modes at 1646 and 1686 cm^{-1}, the G_SC_S(+) and G_SC_S(-) respectively (green and red panels). At 80 ^oC, this is replaced by a lower intensity, broad coupling signal centred around 1668 cm^{-1}, corresponding to the G_S (yellow

panel), associated with single stranded DNA. A loss of base-stacking and presence of G_S both indicate that melting has occurred.

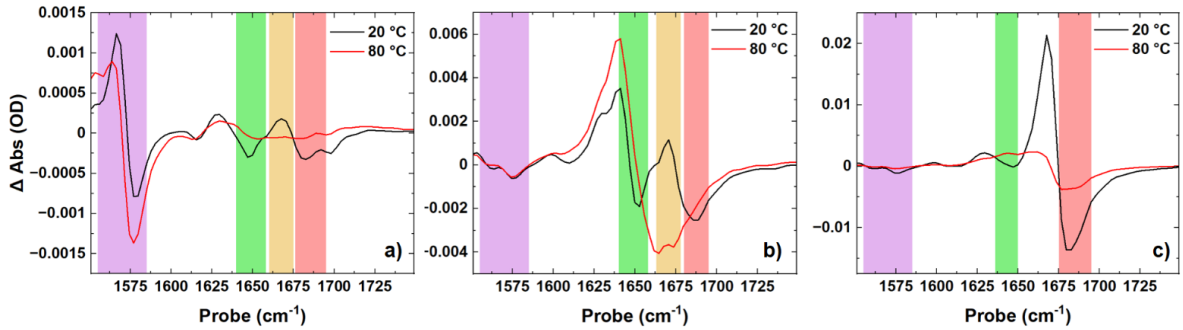

Figure 5.9: *Pump slices of the RNA_{all} 2D-IR taken at 20 °C (black lines) and 80 °C (red lines). Pump slices were taken from the modes that see the greatest change over the melting transition, a) the G_R at 1580 cm^{-1} , b) the G_S at 1668 cm^{-1} and c) the G_SC_S(-) at 1686 cm^{-1}. The vibrational modes are highlighted using coloured panels; the G_R (purple), G_SC_S(+) (green), G_S (yellow) and G_SC_S(-) (red).*

Looking at the pump slices through 1668 cm^{-1} (Fig. 5.9 (b)), at 20 °C there is no G_S mode apparent in this position, which is expected given that the GCs are W-C base paired in the stem under these conditions. However there are two bleach and transient features apparent at 1646 and 1686 cm^{-1}, a result of overlapping of broad G_SC_S(+) and G_SC_S(-) signals in this region. The low temperature coupling observed to the G_R (purple panel) is then also seen due to coupling to these overlapping modes. At 80 °C, the G_SC_S(+) and G_SC_S(-) features are replaced with a single v=0-1 bleach and v=1-2 peak associated with G_S, with only a slight shoulder on the high frequency edge suggesting there might be interference from underlying modes, indicating a high degree of melting. The coupling observed to the G_R does not change over the temperature range. The changes in the G and C modes here are implied in the FT-IR, but the details are obscured by broad overlapping signals, which are much more easily distinguishable in the 2D-IR.

The pump slice at 1686 cm^{-1} (Fig. 5.9 (c)), shows intense peaks for the G_SC_S(-) (red panel) at 20 °C (black trace), which are almost entirely lost on heating to 80 °C (red trace), with some small remnant remaining, corresponding to almost total melting. Cross-peaks are observed at 20 °C, indicative of coupling of the G_SC_S(-) to the G_R and G_SC_S(+) at 1580 and 1646 cm^{-1} respectively, which is also lost on heating.

5.4.3 DNA_{all} 2D-IR

For the DNA_{all} 2D-IR, at 20 °C the G_R at 1575 cm^{-1}, the A_R at 1622 cm^{-1}, G_SC_S(+) at 1646 cm^{-1} and the G_SC_S(-) at 1686 cm^{-1} are also apparent (Fig. 5.10, (c)), similar to RNA_{all}. The G_SC_S(-) appears very broad. At 80 °C no more distinct G_SC_S(+) and G_SC_S(-) are observable, a sign of complete melting. Instead a single broad bleach is seen with a minimum at 1668 cm^{-1}, assigned to the overlapping G_S and C_S.

Focusing on DNA_{all}, the pump slice at 1580 cm^{-1} (Fig. 5.11 (a)) reveals similar coupling for the G_R (purple panel) at 20 °C (black trace) compared to RNA_{all}, with visible cross-peaks to G_SC_S(+) at 1646 cm^{-1}(green panel), but a stronger coupling signal to G_SC_S(-) at 1686 cm^{-1}(red panel). On melting at 80 °C (red trace), a large increase in the G_R is observed with signals three times the size of their 20 °C counterparts and proportionally larger response than observed for RNA_{all}. There is also a small redshift in frequency only observed in DNA_{all}. The coupling of the G_R also changes as expected; with the sequence melted, coupling is seen to the 1668 cm^{-1} G_S (yellow panel), with signal sizes similar to those seen to the paired G modes at low temperature. RNA_{all} sees a much less intense cross-peak on melting, as a result of a smaller increase in the G_R.

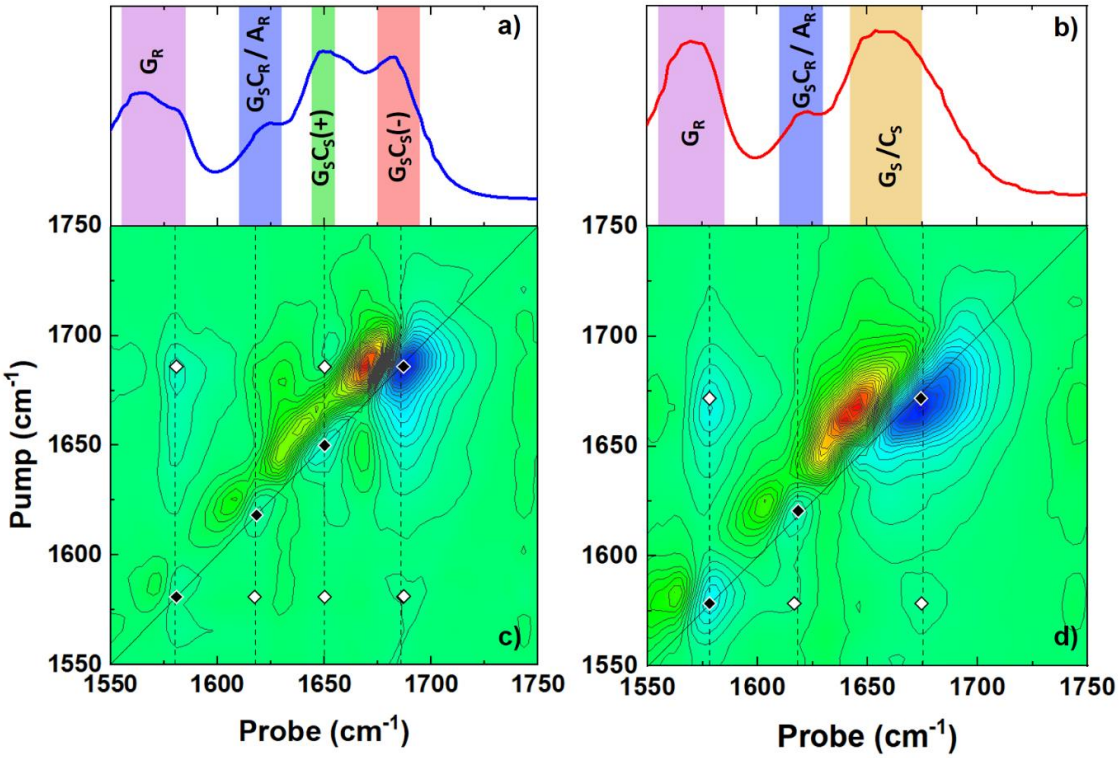

Figure 5.10: *IR spectra showing, FT-IR absorption spectra of DNA_{all} a) at 20 °C and b) 80 °C. Vibrational mode assignments are indicated by coloured panels:, G_R (1575 cm^{-1}, purple); overlapping A_R and G_SC_R (1620 cm^{-1}, blue); G_SC_S(+) (1648 cm^{-1}, green); overlapping G_S, C_S (1656 cm^{-1}, yellow); G_SC_S(-) (1686 cm^{-1}, red).*[39,50,51] *The notation used to identify the vibrational modes of the bases is consistent with that used in Refs 39, 50 & 51 Spectra have been solvent corrected. c) The 20 °C, 250 fs 2D-IR for DNA_{all}. d) The 80 °C, 250 fs 2D-IR for DNA_{all}. Overlaying guidelines indicate the position of modes identified in the IR absorption spectra. On diagonal features are indicated by black diamonds, while cross-peaks are shown by white diamonds.*

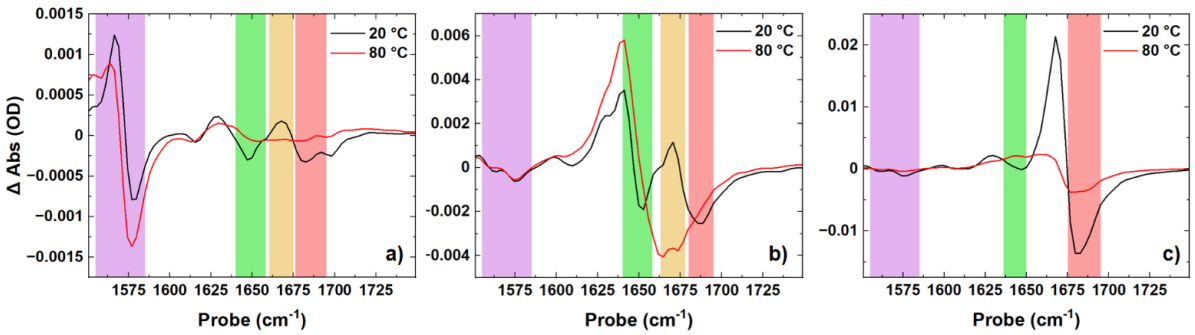

Figure 5.11: *Pump slices of the DNA_{all} 2D-IR take at 20 °C (black) and 80 °C (red). Pump slices were taken from the modes that see the greatest change over the melting transition, a) the G_R at 1580 cm^{-1} , b) the G_S at 1668 cm^{-1} and c) the G_SC_S(-) at 1686 cm^{-1}. The vibrational modes are highlighted using coloured panels; the G_R (purple), G_SC_S(+) (green), G_S (yellow) and G_SC_S(-) (red).*

At 1668 cm^{-1} (Fig. 5.11 (b)), as expected there is no G_S bleach apparent at 20 °C, however there is overlap seen with the G_SC_S(-) of 1686 cm^{-1}. Distortion in the line shape make some interference apparent, likely due to some contribution from the 1646 cm^{-1}, but significantly less than observed in RNA_{all}. There is also coupling to the G_R observed. By 80 °C, there is no trace of the G_SC_S(-), with the bleach shifted to 1668 cm^{-1}, indicating the G_S and no distortion in the line shape indicates full melting. The coupling to the G_R also increased with the rise of the G_S and is seen to shift to lower wavenumber.

With respect to the DNA_{all} pump slices at 1686 cm^{-1} (Fig. 5.11 (c)), strong peaks are seen for the G_SC_S(-) at 20 °C, with clear coupling to the G_R. There is also possible coupling to the G_SC_S(+), though this is somewhat convoluted with the v=1-2 peak of the G_SC_S(-), which appears to be broader than observed in RNA_{all} where the coupling is clear. At 80 °C, there is significant loss of the G_SC_S(-), although some signal remains, though it appears broadened and possibly shifted to lower wavenumber compared to the RNA_{all} response. The coupling to the G_R is also reduced and shifted to lower wavenumber.

Overall, the overarching response in both RNA_{all} and DNA_{all} sees a change in the guanine modes and their coupling over the melt. The guanine vibrational modes G_R, $G_SC_S(+)$ and $G_SC_S(-)$ show clear coupling to each other at low temperature, which shifts at high temperature as $G_SC_S(-)$ is lost during hairpin melting and the G_S mode appears at lower frequency. The 2D-IR also resolves the G_S mode, which cannot be distinguished from the overlapping C_S mode in the IR absorption spectrum.

5.4.4 Spectral Diffusion

The 2D-IR were measured over a range of T_W from 250 fs to 3 ps. This time resolved data enables the observation of spectral diffusion of the inhomogeneously broadened lineshapes, providing insight into the dynamic environment of the modes. To quantify the spectral diffusion of the G_R between high and low temperature, the nodal line slope was determined for the 20 and 80 °C spectra from 250 fs to 3 ps. This shows the dynamics of the sequences in the folded hairpin configuration and single stranded. 2D-IR plots over the waiting times for DNA_{all} are show in Fig. 5.12.

The nodal lines between the v=0-1 and v =1-2 for the G_R have been overlaid (white lines) to illustrate the degree of spectral diffusion over time. At 20 °C (Fig. 5.12, (a-c)), inhomogeneous broadening of the bleach is apparent, but the lineshape is also distorted by coupling to a mode seen on the diagonal at 1560 cm^{-1} (guanine ring C=N). This is still clearly seen at 1 ps. There is also pump scatter at all time points, seen at ~1565 cm^{-1} across the pump axis (most clear in bottom right quadrant in the off diagonal region), which is also seen to cut into the lineshape of the G_R bleach. This is exaggerated by the small signal sizes of the G_R near the edge of the pump profile. At 80 °C (Fig. 5.12, (d-f)), no coupling

is seen, but a broad positive peak is seen in the off diagonal region that overlaps with the G_R v =1-2 peak and may be influencing the nodal line. This feature also persists at late time. In general, the spectral features in this region are seen to influence the lineshapes of neighbouring features, which make analysis of the spectral diffusion challenging. The spectra for RNA_{all} display similar features that distort the lineshapes, though less scatter is observed (See appendix, Fig. A5.1).

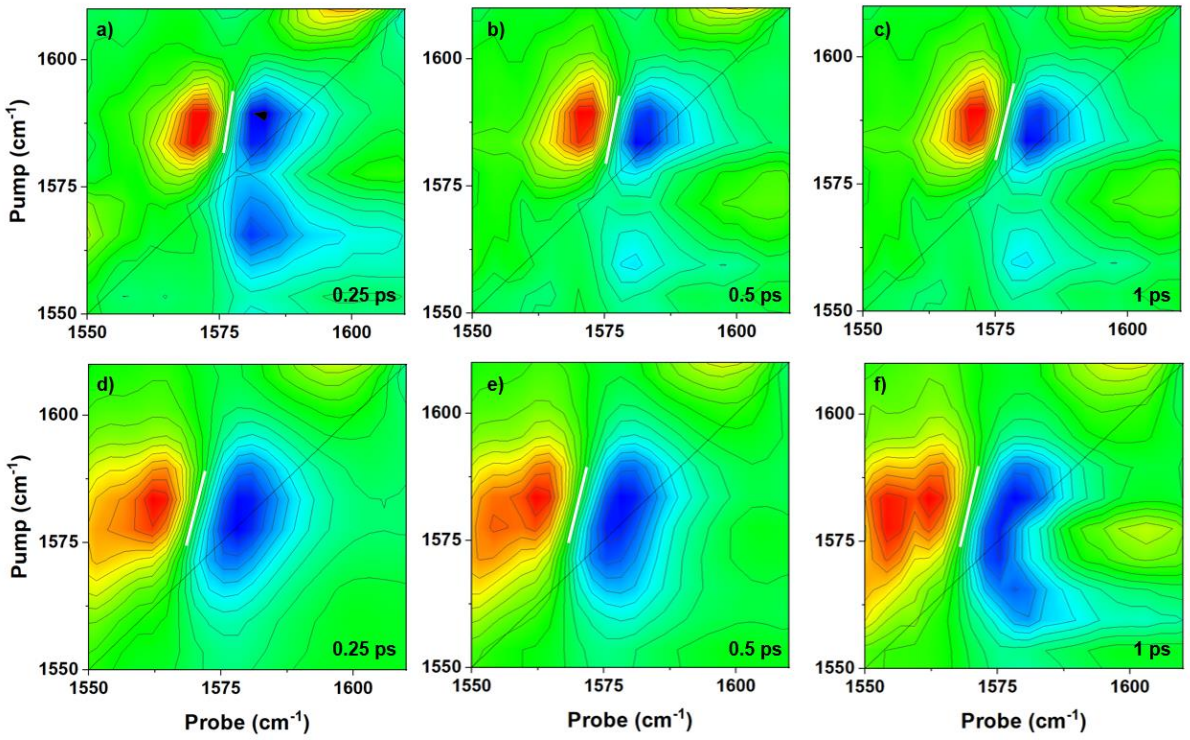

Figure 5.12: *DNA_{all} 2D-IR spectra of the lineshape dependence of the G_R mode on the waiting time and temperature. a-c) shows spectra taken for T_w = 250 fs, 500 fs and 1 ps at 20 °C. d-f) shows spectra taken for T_w = 250 fs, 500 fs and 1 ps at 80 °C. The white lines show the results of a nodal line slope analysis.*

The spectral diffusion as determined by nodal line slope analysis of the G_R modes of RNA_{all} and DNA_{all} over time can be seen in Figure 5.13.[37,38] At low temperature DNA_{all} appears to be more inhomogeneously broadened along the diagonal than RNA_{all}, with relatively slow spectral diffusion though the degree of spectral diffusion in both cases it relatively small. The nodal line slope values for RNA_{all}

are somewhat noisey, but are consistently lower than DNA_{all}, in most cases much lower, though it is not possible to make definite statements due to the uncertainty. At 80 °C, DNA_{all} appears to undergo a greater degree of spectral diffusion than at At 20 °C, though only slightly, and the difference may be within error. Again, RNA_{all} appears to less inhomogeneously broadened, on average having a lower line slope value, but there is a high degree of variability between time points and no clear trend in RNA_{all} can be determined.

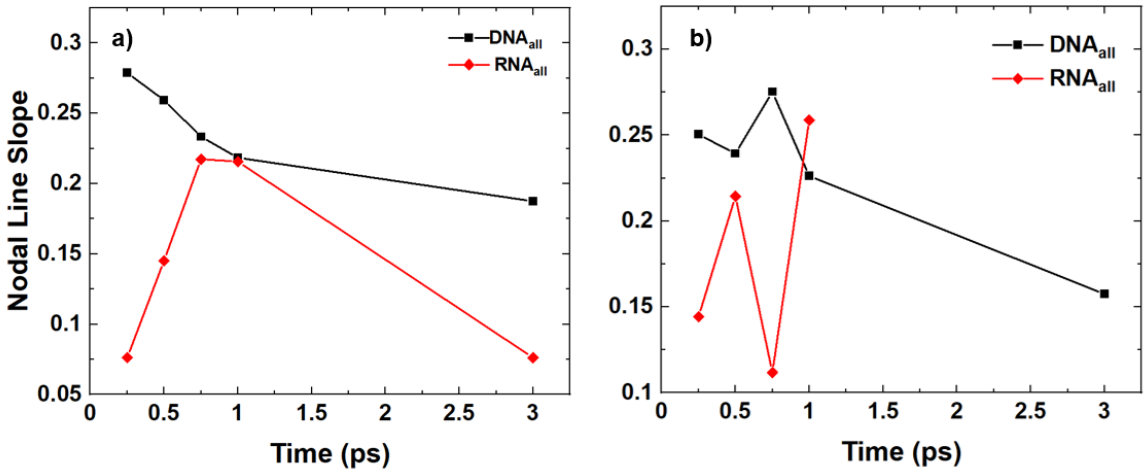

Figure 5.13*: RNA_{all} (red) and DNA_{all} (black) spectral diffusion as determined by the change in nodal sline slope over time at a) 20 °C and b) 80 °C.*

5.4.5 Temperature-Jump 2D-IR Infrared Spectroscopy

With the 2D-IR fully assigned and the static behaviour of the modes and coupling understood between the hairpin and melted single strands of the RNA_{all} and DNA_{all} tetraloops, it enables the characterisation of the intermediate melted spectra. For the Dynamic T-jump 2D-IR measurements, the samples were heated to a T_0 of T_m – 5 °C in order to observe the greatest change between 2D-IR measurements over a fixed T-jump of 10 °C through the melting transition. These spectra were taken at 2D-IR T_w of 250 fs, and T-jump delays times of -20

ns, 500 ns and near the peak melting time, which varies from sequence to sequences with RNA_{all} having a peak melting time of 30 μs, while DNA_{all} has a peak melting time of 5.4 μs (see Chapter 3).

Due to the temperature difference between 2D-IR measurements through the T-jump being equal or less than 10 °C, the changes observed are more subtle than over the entire melt, so difference spectra and difference pump slices will be presented to more clearly highlight the changes seen in the structure. Also peaks in the differenced T-jump 2D-IR are smaller compared to the equilibrium 2D-IR since they are the difference over a fraction of the temperature range, and due to the lower repetition rate of the T-jump 2D-IR as the sample must cool in between T-jump, some data quality is sacrificed as well. As a consequently some of the fine detail and coupling are less clear, which makes the static 2D-IR all the more important as a reference.

5.4.6 RNA_{all} T-jump 2D-IR

2D-IR difference spectra for RNA_{all} are shown in Fig. 5.14, giving a direct comparison between the differences in the equilibrium 2D-IR seen over the full temperature range from 20 – 80 °C (Fig. 5.14 (a)) and the maximum change seen through the T-jump from -20 ns before the T-jump to RNA_{all} peak melting time around 30 μs (Fig. 5.14 (b)).

The clearest features apparent in the equilibrium difference 2D-IR (Fig. 5.14 (a)) are the loss of the 1686 cm^{-1} G_SC_S(-) v=0-1 bleach, signified by a large positive change in the spectrum, and v=1-2 seen as an even more intense negative change in the spectrum, as well as the appearance of a new v=0-1 bleach at 1668 cm^{-1} and associated v=1-2 peak of the G_S, seen as a negative and positive change

in the spectrum respectively in that region. As established in §5.4.2, this loss of coupled W-C modes and appearance of G_S show that melting is occurring. Although much less intense, the negative increase of the v=0-1 and v=1-2 of the G_R at 1580 cm^{-1} can be seen.

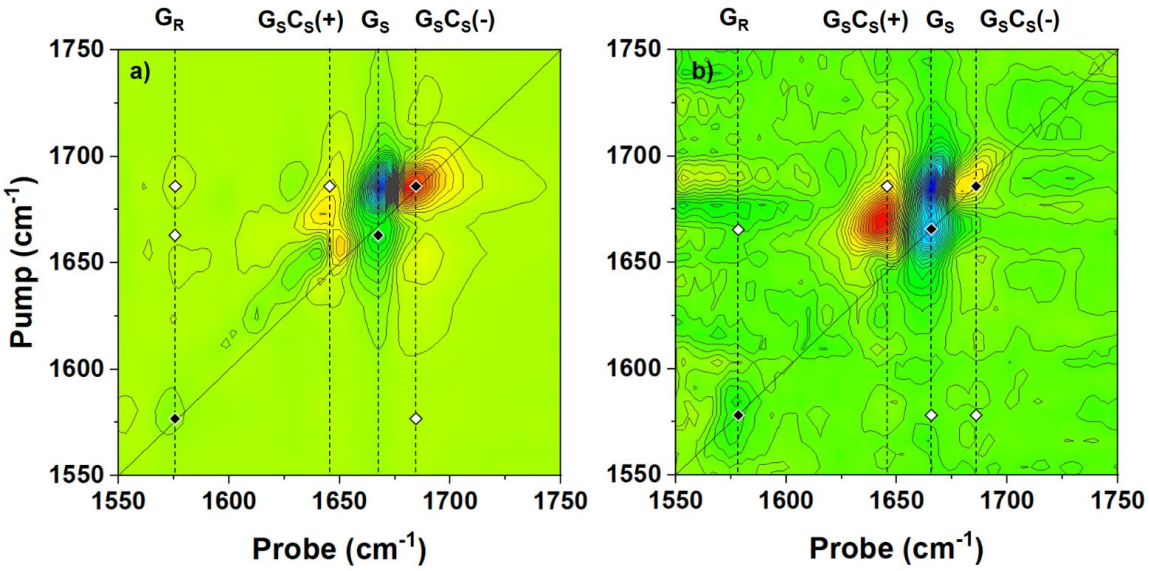

Figure 5.14: *RNA_{all} 2D-IR difference spectra showing a) the equilibrium 80 – 20 °C 250 fs 2D-IR difference spectra, b) the 75 °C T-jump 2D-IR maximum 30 µs – -20 ns 250 fs 2D-IR difference spectra. Changes in on diagonal features are indicated by black diamonds with dashed lines identifying the positions of the G_R (1580 cm^{-1}), G_SC_S(+) (1646 cm^{-1}), G_S (1668 cm^{-1}) and G_SC_S(-) (1686 cm^{-1}). Changes in cross-peaks are indicated by white diamonds.*

The same pattern of changes is also seen in the maximum change observed over the T-jump 2D-IR (Fig. 5.14 (b)), although the signals are less intense. However, this is indicative that melting is being observed in the T-jump 2D-IR.

To determine how changes seen across the T-jump progress during the melting process however, it is necessary to look at time intervals intermediate in the T-jump. For the interval from the T-jump to 500 ns, the differences seen in this interval are capture in the 2D-IR seen in figure 5.15 (a). Changes in the difference

spectra due to melting are clear, with the loss of the transient of the G_SC_S(-) being apparent, although the less intense, broader bleach is not clearly visible on the diagonal. An increase in the G_S peaks is also seen to appear at 1668 cm^{-1}.

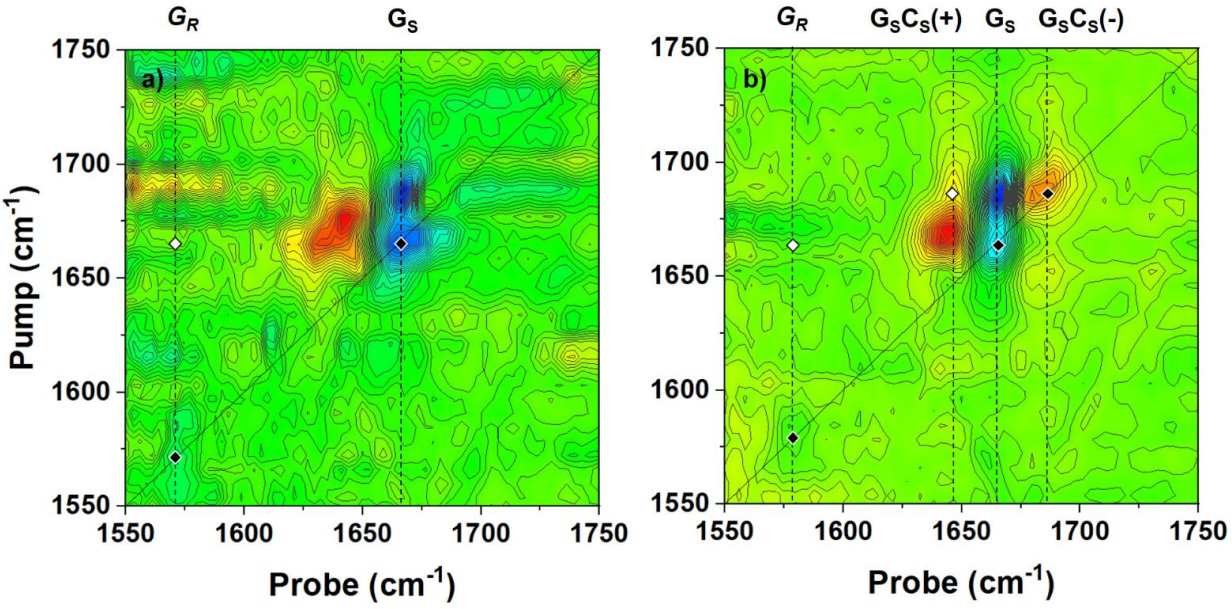

Figure 5.15: *RNA_{all} 2D-IR difference spectra showing a) the 75 °C T-jump 2D-IR early time 500 ns – -20 ns 250 fs difference spectra and b) the 75 °C T-jump 2D-IR late time 30 µs – 500 ns 250 fs difference spectra. Changes in on diagonal features are indicated by black diamonds with dashed lines identifying the positions of the G_R (1580 cm^{-1}), G_SC_S(+) (1646 cm^{-1}), G_S (1668 cm^{-1}) and G_SC_S(-) (1686 cm^{-1}). Changes in cross-peaks are indicated by white diamonds.*

From 500 ns to the maximum difference at 30 µs (Fig. 5.15 (b)), the signals are much stronger, signifying the majority of melting of the sequence occurs in this interval. The peaks of the G_SC_S(-) are now both clearly apparent and even some loss of coupling of the G_SC_S(-) to the G_SC_S(+) (1646 cm^{-1} probe, 1686 cm^{-1} pump) seen as a positive change in the off-diagonal region of spectrum, although it overlaps somewhat with the positive change due to the v= 1-2 of the G_S (1668 cm^{-1}). The G_S too sees the majority of its change after 500 ns with additional

growth in the peaks at 1668 cm^{-1}, although in general the intensity is less than the change seen in the G_SC_S(-).

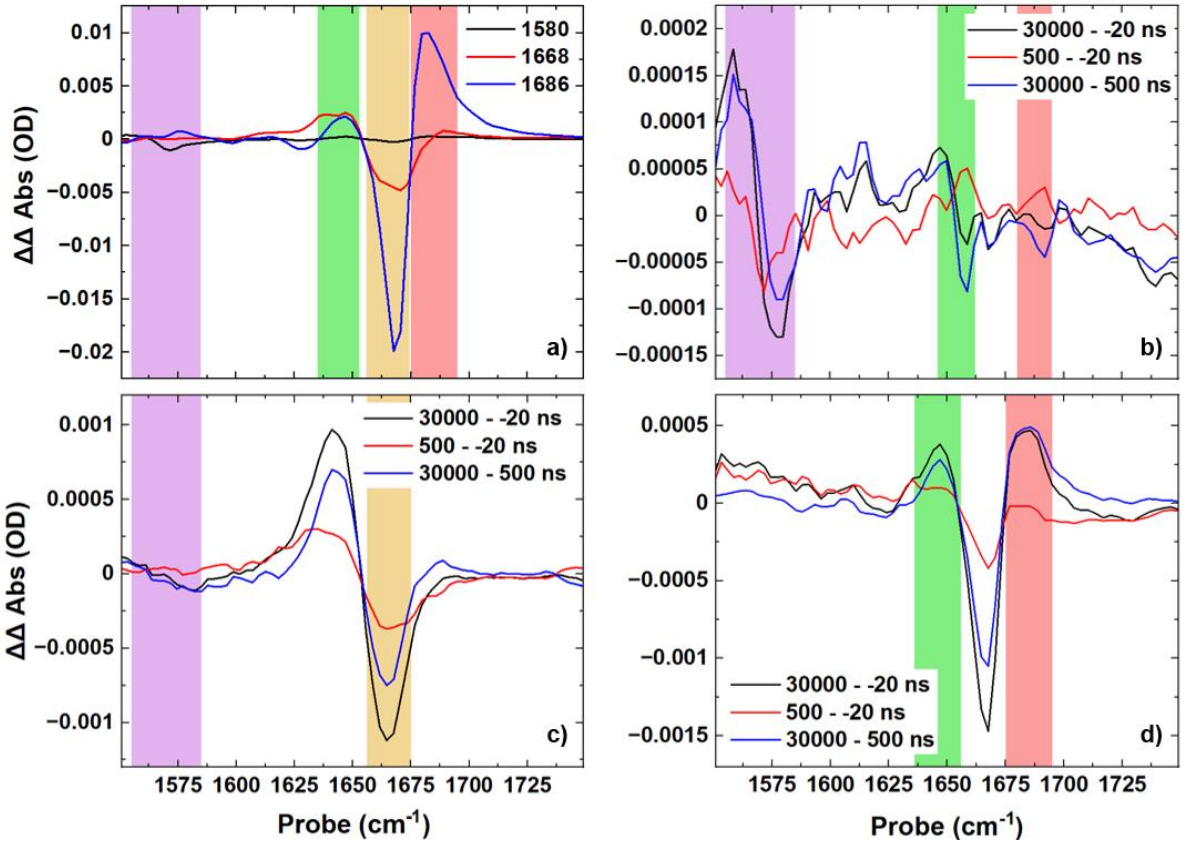

Figure 5.16: *RNA_{all} difference pump slices for a) overlay of the equilibrium 80 – 20 °C 250 fs 2D-IR difference spectra pump slices taken from 1580 cm^{-1} (black), 1668 cm^{-1} (red) and 1686 cm^{-1} (blue), b) the 75 °C T-jump 2D-IR 250 fs difference spectra pump slices at 1580 cm^{-1} for the full melt across the T-jump (black), the first 500 ns (red), and between 500 ns and 30 µs (blue), c) the 75 °C T-jump 2D-IR 250 fs difference spectra pump slices at 1668 cm^{-1} for the full melt across the T-jump (black), the first 500 ns (red), and from 500 ns to 30 µs (blue), d) the 75 °C T-jump 2D-IR 250 fs difference spectra pump slices at 1686 cm^{-1} for the full melt across the T-jump (black), the first 500 ns (red), and between 500 ns and 30 µs (blue). The vibrational modes are highlighted using coloured panels; the G_R (purple), G_SC_S(+) (green), G_S (yellow) and G_SC_S(-) (red).*

Examining the difference pump slices for RNA_{all}, the 2D-IR 80 – 20 °C difference pump slices (Fig. 5.16 (a)) show the changes discuss in detail previously, while also illustrating the relative size of the maximum changes. For instance the change in the 1580 cm^{-1} difference pump slices is a factor of 20 smaller than the change in the G_SC_S(-). This is partly due to the profile of the pump, such that the G_R is on the outer edge of the pump bandwidth. This is clearly seen in the 1580 cm^{-1} signals for the T-jump 2D-IR (Fig. 5.16 (b)), where this region has poor signal to noise as a direct consequence. The main G_R bleach however gains 50% of its intensity before 500 ns (negative change in the bleach). There also appears to be visible coupling, with two distinct cross-peaks consistent with the G_S and G_SC_S(-), although the precise classification is difficult however due contribution of the noise. This suggests that although some melting has clearly occurred, it has not melting fully.

The RNA_{all} T-jump 2D-IR difference pump slices for 1668 cm^{-1} show 30% of the G_S ν=0-1 bleach is gained in first 500 ns (Fig. 5.16 (c)), which is consistent with the G_R in suggesting a not insignificant proportion of melting happens relatively quickly. Coupling to the G_SC_S(+) at 1646 cm^{-1} is not apparent, but an increase in coupling (negative change) is observed to the G_R after 500 ns and is very broad.

Finally, the 1686 cm^{-1} difference pump slice shows the sharp loss (negative change) of the G_SC_S(-) ν= 1-2 peak over the T-jump (Fig. 5.16 (d)), while the loss of the bleach (positive change) is broader and more rounded. In the first 500 ns, a third of the signal is already lost, which is consistent with the G_S. No coupling to the G_SC_S(+) is observable and although a broad change is seen in the region of the G_R at 1580 cm^{-1} from 500 ns to peak melting at 30 μs, this is likely a slight baseline variation that is exaggerated by the small signal.

5.4.7 DNA_{all} T-jump 2D-IR

With respect to the T-jump 2D-IR difference spectra for DNA_{all}, the changes apparent in the 80 – 20 °C difference 2D-IR and the maximum change seen through the T-jump from – 20 ns to DNA_{all} peak melting time around 5.4 µs are shown in Fig. 5.17 (a) and (b).

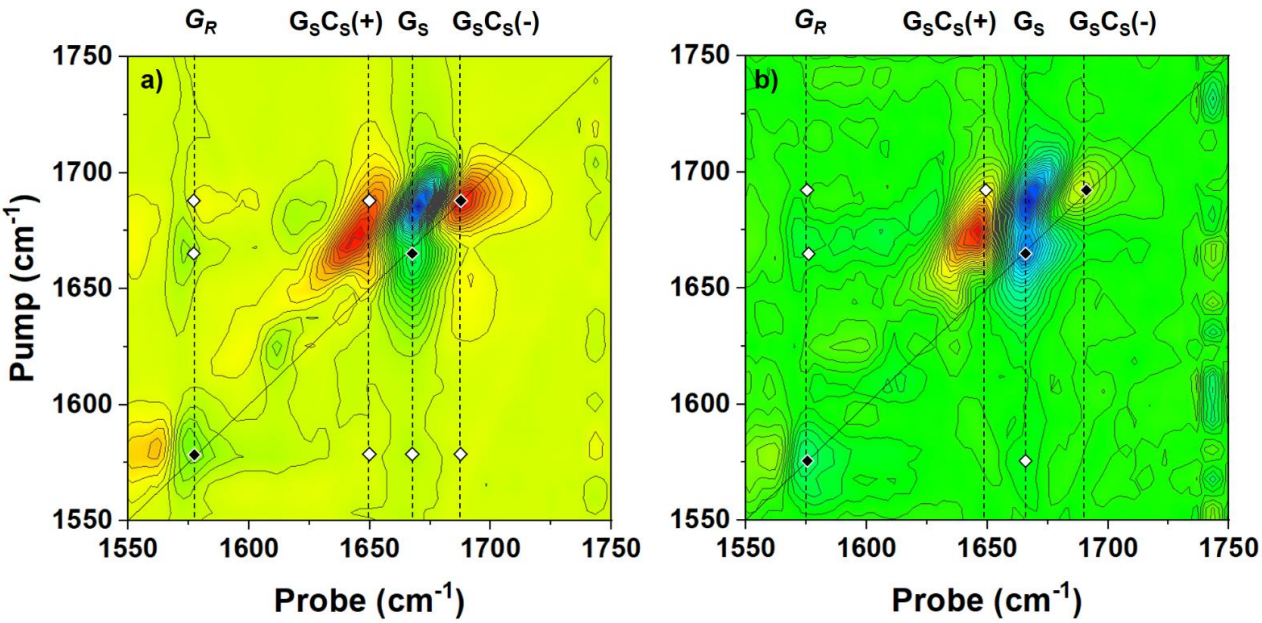

Figure 5.17: *DNA_{all} 2D-IR difference spectra showing a) the equilibrium 80 – 20 °C 250 fs 2D-IR difference spectra, b) the 70°C T-jump 2D-IR maximum 30 µs – -20 ns 250 fs 2D-IR difference spectra. Changes in on diagonal features are indicated by black diamonds with dashed lines identifying the positions of the G_R (1580 cm^{-1}), $G_SC_S(+)$ (1646 cm^{-1}), G_S (1668 cm^{-1}) and $G_SC_S(-)$ (1686 cm^{-1}). Changes in cross-peaks are indicated by white diamonds.*

The same pattern as seen across the equilibrium and maximum T-jump difference 2D-IR in RNA_{all} (Fig, 5.14, (a, b)) are seen in DNA_{all} difference spectra as well, showing the gain of the G_R at 1580 cm^{-1} and G_S at 1668 cm^{-1} and the loss of the $G_SC_S(-)$ at 1686 cm^{-1}

In the fast processes seen in the first 500 ns after the T-jump (Fig 5.18 (a)), the DNA_{all} T-jump 2D-IR difference spectra show a large loss in the 1686 cm^{-1} signal, with a strong positive change in the difference spectrum apparent corresponding to the v=0-1 bleach of the G_SC_S(-) bleach and a similar negative change in the v=1-2 transient. For the G_S however, there is only a small change in the signal, with a slight negative change indicating the growth of the G_S peaks due to melting. This is distinctly different behaviour than was observed in RNA_{all}.

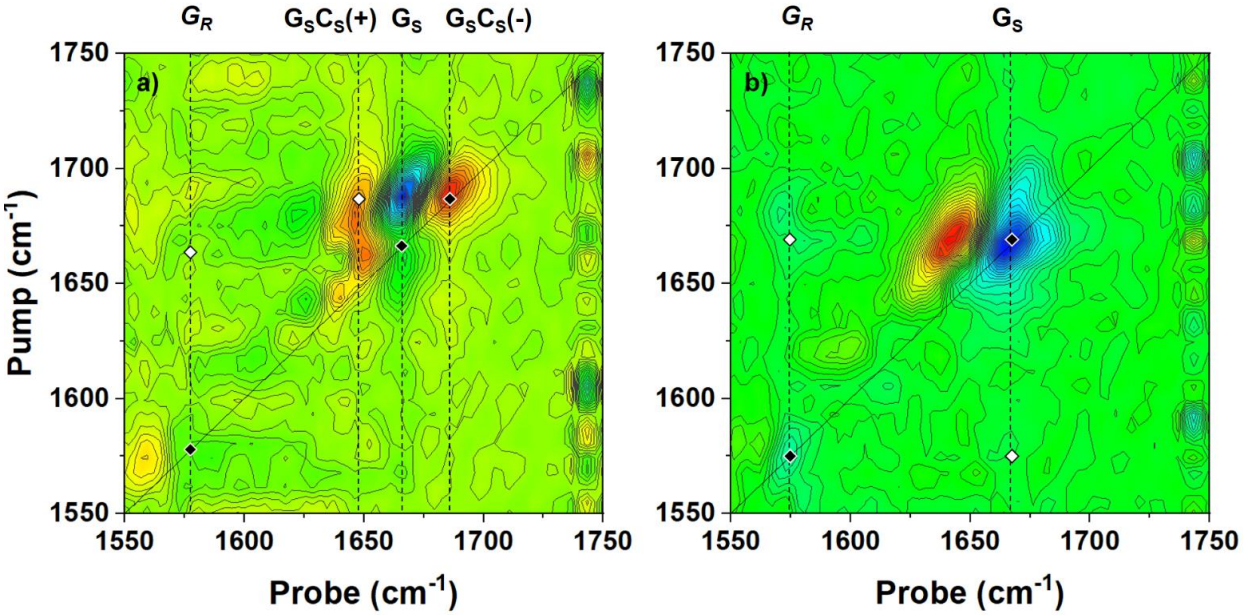

Figure 5.18: *DNA_{all} 2D-IR difference spectra showing a) the 70 °C T-jump 2D-IR early time 500 ns – -20 ns 250 fs difference spectra and b) the 70 °C T-jump 2D-IR late time 30 µs – 500 ns 250 fs difference spectra. Changes in on diagonal features are indicated by black diamonds with dashed lines identifying the positions of the G_R (1580 cm^{-1}), G_SC_S(+) (1646 cm^{-1}), G_S (1668 cm^{-1}) and G_SC_S(-) (1686 cm^{-1}). Changes in cross-peaks are indicated by white diamonds.*

In the late melting phase from 500 ns to 5.4 µs (Fig. 5.18 (b)), the spectrum is dominated by the increase in the G_S peaks, with both an intense and broad change seen in both. By contrast, there is no significant contribution from the

G_SC_S(-), with only a slight shoulder on the high wavenumber top of the G_S bleach to indicate the loss of the G_SC_S(-) transient.

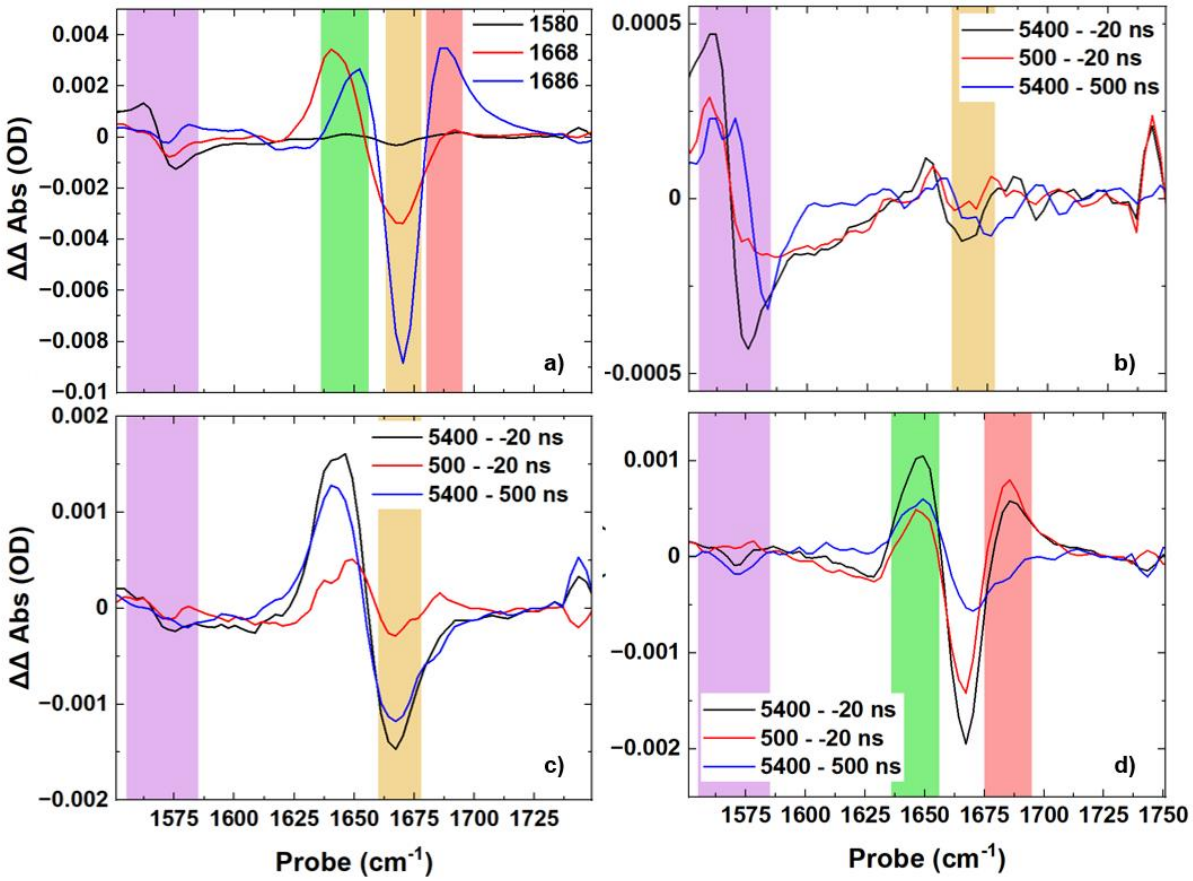

Figure 5.19: *DNA_{all} difference pump slices for a) overlay of the equilibrium 80 – 20 °C 250 fs 2D-IR difference spectra pump slices taken from 1580 cm^{-1} (black), 1668 cm^{-1} (red) and 1686 cm^{-1} (blue), b) the 70°C T-jump 2D-IR 250 fs difference spectra pump slices at 1580 cm^{-1} for the full melt across the T-jump (black), the first 500 ns (red), and between 500 ns and 30 µs (blue), c) the 70°C T-jump 2D-IR 250 fs difference spectra pump slices at 1668 cm^{-1} for the full melt across the T-jump (black), the first 500 ns (red), and from 500 ns to 30 µs (blue), d) the 70 °C T-jump 2D-IR 250 fs difference spectra pump slices at 1686 cm^{-1} for the full melt across the T-jump (black), the first 500 ns (red), and between 500 ns and 30 µs (blue). The vibrational modes are highlighted using coloured panels; the G_R (purple), G_SC_S(+) (green), G_S (yellow) and G_SC_S(-) (red).*

On inspection of the difference pump slices for the 80 – 20 °C DNA_{all} 2D-IR (Fig. 5.19 (a)), the size of the G_R change in the 1580 cm^{-1} pump slice (black line) is conserved compared to RNA_{all}, but the size of the G_SC_S(-) difference is about half. This is likely due to the G_SC_S(-) in general being less intense and broader in DNA.

The 1580 cm^{-1} T-jump 2D-IR difference pump slices (Fig. 5.19 (b)) show an increase in the G_R v= 0-1 bleach in first 500 ns that is very broad, but after 500 ns appears narrow. They contribute similar intensity, with the time after 500 ns account for the greater share at about 60% of the final signal. The region from 1600 to 1700 cm^{-1} is noisy, but coupling to a single broad mode, the G_S appears after 500 ns.

For 1668 cm^{-1} (Fig, 5.19 (c)) only a small gain in the size of the peaks is observed before 500 ns, with the majority of the growth in the signal occurring between 500 ns and 5.4 μs. Similarly, an increase in coupling to the G_R is apparent between 500 ns and 5.4 μs, although the region is noisy.

At 1686 cm^{-1}, the difference pump slices (Fig. 5.19 (d)) show that most of the G_SC_S(-) is lost in the first 500 ns, accounting for ~75% of the signal. The changes after 500 ns however are notable as they show a broadening of the signal. The 1580 cm^{-1} coupling is seen to become more negative after 500 ns, although once more the signals are small enough that the baseline drift between 2D-IR measured at different T-jump delays makes the size somewhat ambiguous.

5.5 Discussion

Both T-jump and 2D-IR are mature technologies in their own right, complementing each other in probing a system in different energy regimes on different time scales; T-jump measuring ground state dynamics occurring on the timescale of nanoseconds to milliseconds, with recent developments to extend that further,[7] while 2D-IR captures the excited state dynamics on the order of femtoseconds to picoseconds. With T-jump 2D-IR, a whole array of new experiments becomes available, providing real time information on the changes in excited state relaxation times, spectral diffusion and coupling during critical stages in a process under investigation that provide insight into the dynamic structural and energetic relationships that are involved.

The use of 2D-IR and T-jump methods to individually interrogate DNA is well established. Despite similarly IR active nucleotide bases comprising RNA, its more varied functional structures and dynamics are much less well studied and understood. By combining the two methodologies and applying it to the hairpin melting process of RNA_{all} and DNA_{all} previously characterised by T-jump IR in Chapter 3,[31] and using DNA_{all} and the depth of literature on DNA 2D-IR in general as a baseline for comparison, it is possible to characterise the behaviour of the RNA UNCG hairpin stem-loop structure.

It is known that in general RNA melts an order of magnitude slower than DNA. This is due to the slower dissociation of the base-stacking as a result of the ordered hydration of the backbone and tighter packing of the bases.[33,57–59] From previous work it was determined that specifically RNA_{all} has a lifetime of melting of 6 μs at 70 °C, compared to DNA_{all} which has a lifetime of melting of 0.8 μs under the same conditions.[31]

5.5.1 RNA_{all} Hairpin

In the RNA_{all} T-jump 2D-IR, 50% of the total change in the G_R signal is seen in the first 500 ns. This compares with 30% of the total G_S and a corresponding 30% loss of the G_SC_S(-). This change represents a larger proportion of the signal than seen in the 1D T-jump (see Chapter 3). All of these modes relate to the environment of the guanine bases, their base-stacking and base-pairing interactions, so it is interesting to note that the proportion of the G_R modes not relate to that of the G_SC_s(-) or the G_S mode.

The melting of the stem base pairs consists primarily of two processes, the breakdown of base–stacking and base-pairing. The G_R mode is suppressed by base stacking so it primarily, though not exclusively, reports on the state of base stacking in the stem. The G_SC_S(-) mode is produced by the coupling of the carbonyl stretching modes of the guanine and cytosine bases that is possible when the two bases are hydrogen bonded, forming a W-C base pair. As a consequence the mode is lost when the GC base pairs in the stem dissociate, and its presence is a reliable indicator of the presence of base-pairing. However, it is not necessarily true that the base-stacking has no influence on the mode.

In the time between 500 ns and 30 μs when the maximum melting signal for RNA_{all} is observed, the rest of the melting occurs. Over the T-jump, changes appear in the signal size of the modes, with no shifts in wavenumber observed. Due to the small signal sizes and noise, the precise determination of the behaviour of the coupling is uncertain between separate T-jump 2D-IR measurements, but at late time there still appears to be coupling between the G_R and G_SC_S(-), suggesting that although the G_SC_S(-) has reduced in size, it is still present, but an increase in coupling between the G_R and G_S is observed. The G_S gains 70% of its signal in this interval, with a commensurate loss seen in the

G_SC_S(-), showing that melting is occurring. However it clearly does not result in full dissociation, with some population of the hairpins showing signs of base-pairing. This is consistent with the melting curves for RNA_{all} which were not fully melted even at 95 °C (Ch. 3, §3.3.1), due to having a high melting point and a relatively broad melting range.

5.5.2 DNA_{all} Hairpin

In the DNA_{all} tetraloop T-jump 2D-IR, the early time response before 500 ns accounts for roughly 40% in the increase of the G_R mode and is a very broad response. This compares to the 50% seen in the 1D T-jump (See Chapter 3). The G_SC_S(-) however see a loss of greater than 75% in the same time interval, while the G_S only grows by about 25% of the total chance seen over the T-jump in this time. Additionally calculations conducted by Cho and co-workers determined that the G_S blue-shifts by up to 50 cm^{-1} (over the full melt).[50] This suggests that the loss of W-C coupled modes and ones associated with single stranded sequence do not necessarily reflect the same degree of change simultaneously and one may respond more quickly to a change in the structure than another under certain conditions.

The response of the G_R from 500 ns to the peak melting time, is much sharper than in the first 500 ns. This also hints at a difference in the conformational dynamics between the two time intervals in the melting process. The G_SC_S(-) sees the opposite trend, with the late time loss of the signal being much more broad. The coupling is noisy but is seen to grow at late time between the G_R and the G_S, and strangely it also appears to grow for the G_SC_S(-), even as the mode is being lost. This is counter to expectation, although even at its greatest, the

coupling is not particularly prominent. Comparing to the equilibrium 2D-IR it is noted a similar feature is also visible in the 1686 cm^{-1} difference pump slice.

5.5.3 RNA-DNA Comparison

In general there are much greater distinctions between the behaviour of the different modes in the two time intervals seen in DNA_{all} when compared to RNA_{all}. In $RNA_{all,}$ the $G_SC_S(-)$ and the G_S change in magnitude by the same amount in each time interval. DNA_{all} however sees a significant temporal separation of the change in G_S and $G_SC_S(-)$, with the loss of $G_SC_S(-)$ proceeding much faster than rise in the G_S. These modes can be viewed as two ends of a continuum reporting on the degree of W-C pairing. Therefore this separation may indicate a difference in the melting process between DNA_{all} and RNA_{all}. It's possible that the less flexible RNA structural features and helical conformation that are responsible for the much slower melting dynamics in RNA,[33,60] also limit the range of transitional conformations it can adopt. It should be noted that while the previous studies of RNA_{all} and DNA_{all} focused primarily on comparing the behaviour of the G_R,[31] ostensibly this should be able to be detected in 1D T-jump. However, T-jump 2D-IR has a number of advantages that make this distinction more apparent; the narrower linewidths in the 2D spectrum makes the contributions at frequency much clearer and that due to the difference in their extinction coefficient the contribution from the D_2O dynamics is suppressed, which is not trivial to correct for in conventional T-jump.

Over the course of the T-jump, it also appears as if the $G_SC_S(-)$ of DNA_{all} broadens at later time, an effect that is not seen in RNA_{all}. This is possibly due to the different range of conformations available to DNA compared to the less flexible RNA and therefore results in a more inhomogeneous lineshape as there are

more possible intermediate environments of the GC vibrational modes. This too would suggest that RNA melting pathways are comparatively conformationally limited, while DNA explores the potential energy surface in a more multistate melting transition.

5.6 Conclusions

Using a temperature-jump two-dimensional infrared spectroscopy method it was possible to capture transient 2D-IR spectra at discrete time points during a dynamic melting process of nucleic acids on the timescale of nanoseconds to milliseconds. With this it was possible to track the dynamics and coupling of different GC base modes in analogous RNA UNCG and DNA TNCG tetraloop hairpins.

The dynamics of the different modes in RNA hairpin appeared to change in concert, with no differences in proportion of change observed in the different modes sensitive to W-C base-pairing at the different time points. In DNA however, there was clear temporal separation, with a significant lag in the appearance of the uncoupled mode after the significant loss of the coupled mode was observed and broadening observed late in the melt. This suggested that while RNA is known to melt significantly slower, there is a more concerted dissociation of both the base-stacking and base pairing with the relative inflexibility of the RNA structure resulting in a more sharply delineated melting process compared to DNA that appears to follow a comparatively more multi-state transition.

5.7 Appendix

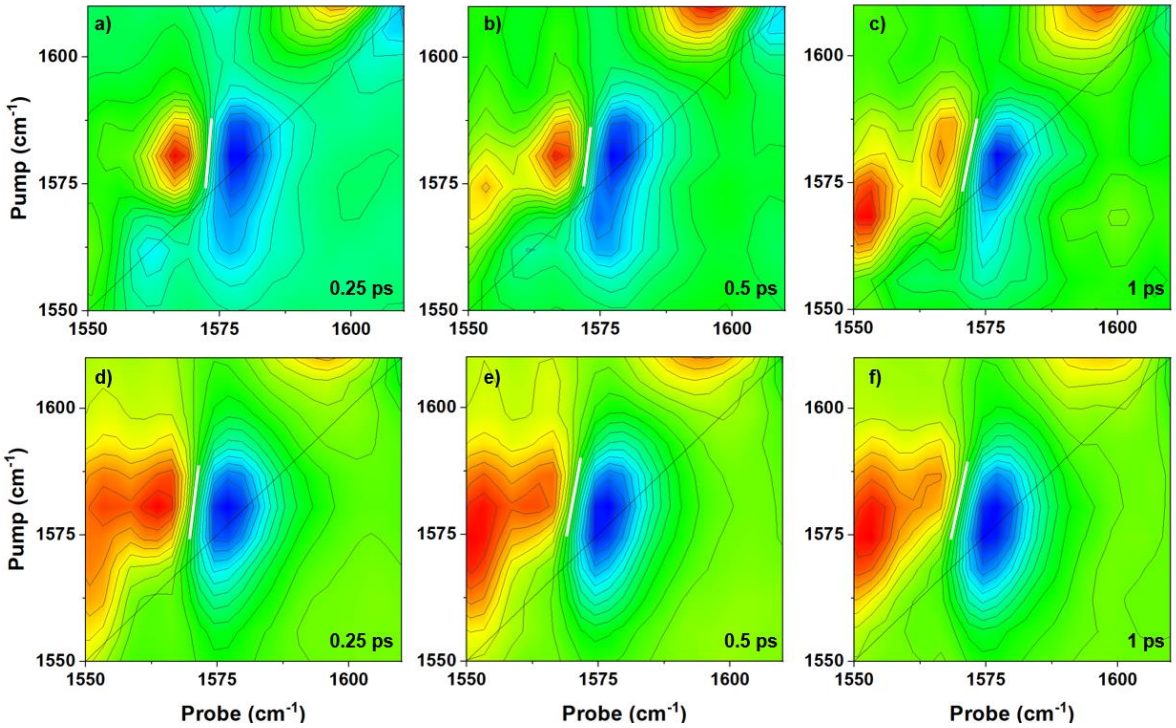

Figure A5.1*: RNA_{all} 2D-IR spectra of the G_R mode of the waiting time and temperature. The top row shows spectra taken for T_w = 250 fs, 500 fs and 1 ps at 20 °C. The bottom row shows the same waiting times at 80 °C. The white shows the results of a nodal line slope.*

5.8 References

1 R. Fritzsch, S. Hume, L. Minnes, M. J. Baker, G. A. Burley and N. T. Hunt, *Analyst*, 2020, **145**, 2014–2024.

2 A. T. Krummel and M. T. Zanni, *J. Phys. Chem. B*, 2006, **110**, 13991–14000.

3 A. T. Krummel, P. Mukherjee and M. T. Zanni, *J. Phys. Chem. B*, 2003, **107**, 9165–9169.

4 R. J. Fick, A. Y. Liu, F. Nussbaumer, C. Kreutz, A. Rangadurai, Y. Xu, R. D.

Sommer, H. Shi, S. Scheiner and A. L. Stelling, *J. Phys. Chem. B*, 2021, **125**, 7613–7627.

5 D. A. Price, P. Wedamulla, T. D. Hill, T. M. Loth and S. D. Moran, *Spectrochim. Acta - Part A Mol. Biomol. Spectrosc.*, 2022, **267**, 120596.

6 C. S. Peng, K. C. Jones and A. Tokmakoff, *J. Am. Chem. Soc.*, 2011, **133**, 15650–15660.

7 B. Ashwood, N. H. C. Lewis, P. J. Sanstead and A. Tokmakoff, *J. Phys. Chem. B*, 2020, **124**, 8665–8677.

8 R. Fritzsch, P. M. Donaldson, G. M. Greetham, M. Towrie, A. W. Parker, M. J. Baker and N. T. Hunt, *Anal. Chem.*, 2018, **90**, 2732–2740.

9 G. Hithell, D. J. Shaw, P. M. Donaldson, G. M. Greetham, M. Towrie, G. A. Burley, A. W. Parker and N. T. Hunt, *J. Phys. Chem. B*, 2016, **120**, 4009–4018.

10 L. A. I. Ramakers, G. Hithell, J. J. May, G. M. Greetham, P. M. Donaldson, M. Towrie, A. W. Parker, G. A. Burley and N. T. Hunt, *J. Phys. Chem. B*, 2017, **121**, 1295–1303.

11 X. X. Zhang, S. L. Brantley, S. A. Corcelli and A. Tokmakoff, *Commun. Biol.*, 2020, **3**, 1–9.

12 G. Hithell, G. A. Burley, P. M. Donaldson, N. T. Hunt, A. W. Parker, M. Towrie, K. Wynne, G. M. Greetham and M. González-Jiménez, *Phys. Chem. Chem. Phys.*, 2017, **19**, 10333–10342.

13 G. Hithell, P. M. Donaldson, G. M. Greetham, M. Towrie, A. W. Parker, G. A. Burley and N. T. Hunt, *Chem. Phys.*, 2018, **512**, 154–164.

14 P. J. Sanstead and A. Tokmakoff, *J. Phys. Chem. B*, 2018, **122**, 3088–3100.

15 E. M. Bruening, J. Schauss, T. Siebert, B. P. Fingerhut and T. Elsaesser, *J. Phys. Chem. Lett.*, 2018, **9**, 583–587.

16 B. P. Fingerhut, A. Kundu, J. Schauss and T. Elsaesser, in *Optics InfoBase Conference Papers*, 2020.

17 J. Schauss, A. Kundu, B. P. Fingerhut and T. Elsaesser, *J. Phys. Chem. Lett.*, 2019, **10**, 6281–6286.

18 J. Schauss, F. Dahms, B. P. Fingerhut and T. Elsaesser, *J. Phys. Chem. Lett.*, 2019, **10**, 238–243.

19 J. Schauss, A. Kundu, B. P. Fingerhut and T. Elsaesser, *J. Phys. Chem. B*, 2021, **125**, 740–747.

20 B. P. Fingerhut and T. Elsaesser, *Springer Ser. Opt. Sci.*, 2019, **226**, 171–195.

21 J. Bredenbeck, J. Helbing and P. Hamm, *J. Am. Chem. Soc.*, 2004, **126**, 990–991.

22 N. T. Hunt, *Dalt. Trans.*, 2014, **43**, 17578–17589.

23 J. Bredenbeck, J. Helbing, C. Kolano and P. Hamm, *ChemPhysChem*, 2007, **8**, 1747–1756.

24 J. M. Anna, C. R. Baiz, M. R. Ross, R. Mccanne and K. J. Kubarych, *Int. Rev. Phys. Chem.*, 2012, **31**, 367–419.

25 H. S. Chung, M. Khalil, A. W. Smith and A. Tokmakoff, *Rev. Sci. Instrum.*, 2007, **78**, 063101.

26 N. T. Hunt, *Chem. Soc. Rev.*, 2009, **38**, 1837–1848.

27 X. X. Zhang, K. C. Jones, A. Fitzpatrick, C. S. Peng, C. J. Feng, C. R. Baiz and

A. Tokmakoff, *J. Phys. Chem. B*, 2016, **120**, 5134–5145.

28 M. S. Jones, B. Ashwood, A. Tokmakoff and A. L. Ferguson, *J. Am. Chem. Soc*, 2021, **143**, 17395–17411.

29 B. Ashwood, P. J. Sanstead, Q. Dai, C. He and A. Tokmakoff, *J. Phys. Chem*, 2019, **2020**, 640.

30 P. J. Sanstead, B. Ashwood, Q. Dai, C. He and A. Tokmakoff, *J. Phys. Chem. B*, 2020, **124**, 1160–1174.

31 C. P. Howe, G. M. Greetham, B. Procacci, A. W. Parker and N. T. Hunt, *J. Phys. Chem. Lett.*, 2022, **13**, 9171–9176.

32 J. E. Johnson and C. G. Hoogstraten, *J. Am. Chem. Soc.*, 2008, **130**, 16757–16769.

33 I. Anosova, E. A. Kowal, M. R. Dunn, J. C. Chaput, W. D. V. Horn and M. Egli, *Nucleic Acids Res.*, 2016, **44**, 1007–1021.

34 A. Kundu, J. Schauss, B. P. Fingerhut and T. Elsaesser, *J. Phys. Chem. B*, 2020, **124**, 2132–2138.

35 P. Hamm and M. T. Zanni, *Concepts and Methods of 2D Infrared Spectroscopy*, Cambridge University Press, Cambridge, 2011.

36 K. Kwak, S. Park, I. J. Finkelstein and M. D. Fayer, *J. Chem. Phys.*, 2007, **127**, 124503.

37 A. Ghosh and R. M. Hochstrasser, *Chem. Phys.*, 2011, **390**, 1–13.

38 K. Kwac and M. Cho, *J. Chem. Phys*, 2003, **119**, 2256.

39 M. Banyay, M. Sarkar and A. Gräslund, *Biophys. Chem.*, 2003, **104**, 477–488.

40 R. Fritzsch, G. M. Greetham, I. P. Clark, L. Minnes, M. Towrie, A. W. Parker and N. T. Hunt, *J. Phys. Chem. B*, 2019, **123**, 6188–6199.

41 G. M. Greetham, I. P. Clark, B. Young, R. Fritsch, L. Minnes, N. T. Hunt and M. Towrie, *Appl. Spectrosc.*, 2020, **74**, 720–727.

42 L. Minnes, G. M. Greetham, D. J. Shaw, I. P. Clark, R. Fritzsch, M. Towrie, A. W. Parker, A. J. Henry, R. J. Taylor and N. T. Hunt, *J. Phys. Chem. B*, 2019, **123**, 8733–8739.

43 G. M. Greetham, P. Burgos, C. Qian, I. P. Clark, P. S. Codd, R. C. Farrow, M. W. George, M. Kogimtzis, P. Matousek, A. W. Parker, M. R. Pollard, D. A. Robinson, X. Zhi-Jun and M. Towrie, *Appl. Spectrosc.*, 2010, **64**, 1311–1319.

44 G. M. Greetham, P. Matousek, A. W. Parker and M. Towrie, *Cent. Laser Facil. Annu. Rep.*, 2008, 249–250.

45 S. H. Shim, D. B. Strasfeld, Y. L. Ling and M. T. Zanni, *Proc. Natl. Acad. Sci. U. S. A.*, 2007, **104**, 14197–14202.

46 S. L. D. Wrathall, B. Procacci, M. Horch, E. Saxton, C. Furlan, J. Walton, Y. Rippers, J. N. Blaza, G. M. Greetham, M. Towrie, A. W. Parker, J. Lynam, A. Parkin and N. T. Hunt, *Phys. Chem. Chem. Phys*, 2022, **24**, 24767.

47 E. B. Dunkelberger, M. Grechko and M. T. Zanni, *J. Phys. Chem. B*, 2015, **119**, 14065–14075.

48 S. Hume, G. Hithell, G. M. Greetham, P. M. Donaldson, M. Towrie, A. W. Parker, M. J. Baker and N. T. Hunt, *Chem. Sci.*, 2019, **10**, 6448–6456.

49 M. Khalil, N. Demirdöven and A. Tokmakoff, *J. Phys. Chem. A*, 2003, **107**, 5258–5279.

50 C. Lee, K. H. Park and M. Cho, *J. Chem. Phys.*, 2006, **125**, 114508.

51 C. Lee and M. Cho, *J. Chem. Phys.*, 2006, **125**, 114509.

52 P. J. Sanstead, P. Stevenson, A. Tokmakoff and A. Tokmako, *J. Am. Chem. Soc.*, 2016, **138**, 11792–11801.

53 J. Dale, C. P. Howe, H. Toncrova, R. Fritzsch, G. M. Greetham, I. P. Clark, M. Towrie, A. W. Parker, T. C. McLeish and N. T. Hunt, *Phys. Chem. Chem. Phys.*, 2021, **23**, 15352–15363.

54 C. Lee, K.-H. Park, J.-A. Kim, S. Hahn and M. Cho, *J. Chem. Phys.*, 2006, **125**, 114510.

55 C. Lee and M. Cho, *J. Chem. Phys.*, 2007, **126**, 145102.

56 L. Dascenzo, F. Leonarski, Q. Vicens and P. Auffinger, *RNA*, 2017, **23**, 259–269.

57 W. Zhang and S. J. Chen, *Biophys. J.*, 2006, **90**, 765–777.

58 K. Sarkar, K. Meister, A. Sethi and M. Gruebele, *Biophys. J.*, 2009, **97**, 1418–1427.

59 K. Sarkar, D. A. Nguyen and M. Gruebele, *RNA*, 2010, **16**, 2427–2434.

60 A. Pérez, A. Noy, F. Lankas, F. J. Luque and M. Orozco, *Nucleic Acids Res.*, 2004, **32**, 6144–6151.

6. Conclusions and Future Work

The work presented over the preceding chapters has demonstrated the use of T-jump as a label-free tool to examine nucleic acid melting and refolding dynamics over a timescale of nanoseconds to milliseconds. In Chapter 3 it was shown that a method could be implemented to produce solvent cooling times that were faster than the refolding rate of a short RNA UNCG hairpin sequence and an analogous DNA hairpin. This made it possible to capture their native dynamics for both melting and refolding processes. Comparing the melting of RNA and DNA it was found that the process of strand dissociation takes an order of magnitude longer than observed in DNA, due to differences in their conformations, with the closer base-stacking of the RNA A-form helix providing a greater barrier to strand melting. The rate limiting step in refolding however was found to be the same in both; the formation of the initial correctly positioned base-stack prior to full “zipping” of the sequence. This refolding was shown to proceed via a complex, multistate transition, involving a “rugged” potential energy landscape of partially paired or misfolded intermediate states, resulting in a sizable loss in entropy associated with folding.

In Chapter 4, the sequence dependence of the RNA and DNA hairpin dynamics was explored by introduction of an AT base pair in different positions in the stem, in order to provide a site-specific probe of dynamics. This allowed the comparative melting and refolding pathways to be determined. In DNA it was demonstrated that melting was initiated from the terminal end, with the stem unzipping toward the loop. In RNA however the overall melting times were not seen to be particularly dependent on base sequence, but displayed a melting pathway starting from the loop and progressing though a bubble conformation before complete strand dissociation.

On refolding, the inclusion of an AT was generally seen to slow the rate of association, compared to an all GC stem, suggesting disruption of the folding pathways for both RNA and DNA. But as with melting, RNA and DNA displayed different dynamic behaviour, with the terminal end position being shown to be a critical nucleation point in the refolding process in RNA, before zipping up toward the loop. However, DNA favoured a refolding pathway starting in the loop closing position and zipping toward the end of the stem. This is due to the greater flexibility of the DNA structure, allowing different conformational structures during the refolding process compared to RNA.

The work presented in Chapter 5 focussed on investigating of the vibrational modes of the RNA and DNA tetraloops as a probe of changes in structure during the melting process with a combination of equilibrium 2D-IR to assess the changes in the structural and coupling relationships through strand melting and using a T-jump 2D-IR method making it possible to capture these changes over discrete times intervals during the melting process. This improved on 1D methods by allowed the spectral changes over the melt to be attributed to specific vibrational modes in the complex base region of the nucleic acid IR spectrum. Observation of the RNA tetraloops dynamics through the melting process showed a concerted change in the modes associated with melting, with similar degrees of melting apparent observed in the different modes over each interval. In DNA however, a temporal separation in the appearance of uncoupled modes after significant loss of coupled modes was observed, suggesting a more multistate transition, attributed to the greater flexibility of its structure compared to RNA. This flexibility allows a greater number of transitional conformations to be explored over the melting process.

The T-jump method has been shown to be effective as a label-free method for capturing the dynamics of processes over nanosecond to millisecond timescale. Due to the cooling for a sample varying with the sample pathlength, this method can be tuned to match the dynamic timescale of different processes as appropriate for analysing different systems. Just in its ability to capture the melting and refolding of nucleic acids however there is a lot of potential for further investigation of conformational dynamics associated with different functional structures adopted by RNA and DNA. It has already been demonstrated in this thesis that refolding in short DNA and RNA hairpin changes with sequences, but this may rely on a number of other factors as well that have yet to be explored. For instance, while the terminal end of the four base pair stem was the favoured refolding nucleation site for RNA, this may not be true in longer sequences. By exploring the differences in refolding pathways as a function of sequences length, insight could be gained on the folding and function of larger, self-assembling RNA structures, which could be useful in the development of RNA based nanotechnologies. Changes in dynamics of DNA and RNA can also be investigated in the presence of relevant small biological molecules, in order to gain insight into *in vivo* interactions which are part of their biological functions.

T-jump 2D-IR methods are still relatively new, but they have a great potential for deepening our understanding of conformational dynamics in nucleic acids as well as . With the timescale of a process having been determined through T-jump or other methods, T-jump 2D-IR can be used to probe particular time points of interest associated with key intermediate steps to gain greater insight into complex processes. Along with coupling interactions, the changes in vibrational dynamics, relaxation times and spectral diffusion can also be investigated at critical moments of conformational change.

T-jump IR methods are useful for systems that undergo temperature dependent dynamic processes that result in a spectral changes in the IR region. However, one limitation is the need to use deuterated water for nucleic acids and proteins to observe their spectral features in the 1650 cm^{-1} region where water absorbs strongly. This can influence dynamics and may not be appropriate for all systems. Equally, many biological structures undergo irreversible thermal denaturation, making T-jump methods ill-suited to probe their conformational dynamics.

By enabling the capture of the conformational dynamics of refolding with T-jump IR and providing greater structure insight with 2D-IR methods, it is possible the body of science for nucleic acids. This is critical for improving on the current theoretical and practical understanding of DNA and RNA, allowing the refinement of nucleic acid force-fields to provide better prediction of structure, folding and dynamics and for a more sophisticated, multistate approach to nuclei acid processes, and enabling more targeted development of new biomedical technologies as well.