Effects of High Pressure on DNA and its Components

A thesis presented in partial fulfilment of the requirements for the degree of

PhD
in
Bio Physics

at Massey University, Manawatū,
New Zealand.

Christopher Paul Lepper

Abstract

There have been many speculations for the environment in which life originated but it has still yet to be determined what environmental chemical and physical conditions were necessary for the evolution of self-replicating chemical systems. While it has been determined that DNA, RNA and their components are chemically unstable at high temperatures, there has currently been only a small number of studies into the role of high pressures on the chemical and physical stabilities.

High-pressure NMR spectroscopy has been used here to study the effects of high temperatures/pressures on the chemical stability of DNA and its components. This has been done with the use of a specialised commercial high-pressure NMR cell capable of withstanding pressures up to 250 MPa. In addition to this, a custom safe handling apparatus and pump system was developed for the operation of this cell.

Studies into the effects of high pressures on the rate of hydrolysis of cytosine and cytidine at 100 °C were performed by measuring the rates of hydrolysis with time under various pressure conditions. These results have shown that the rates of hydrolysis of cytosine and cytidine increase considerably with pressure.

The effects of high pressure on the physical stability of DNA were determined by performing dissociation (melting) experiments on several different DNA sequences under multiple pressure conditions. It was found that the melting point of a small DNA hexamer decreased slightly with pressure whereas the melting points of larger dodecamers increased overall with pressure. It was also found that the melting point of an i-motif structure decreased with increasing pressure.

The effects of high pressure on the chemical stability of cytosine were again studied, this time for cytosine residues within both single- and double-stranded DNA. DNA samples for bacteriophage ΦX174 were incubated under various temperature and pressure conditions. Results for these studies have yet to be determined as the incubated DNA is yet to be sequenced.

It has been discovered that high pressures have a negative effect on the chemical stability of DNA constituents while having an overall small positive effect on the physical stability of DNA.

Acknowledgements

Firstly I would like to thank my supervisors, Prof. Geoffrey Jameson, Dr. Patrick Edwards and Prof. Martin (Bill) Williams, for their fantastic help and support throughout this project. I would like to offer a special thanks to Dr. Vyacheslav Filichev, Dr. Mark Patchett and Assoc. Prof. Murray Cox; this project would have not been possible without their expertise and help. I would also like to thank the Massey University IFS Engineering work shop who fabricated a large portion of the equipment for the pump system used in this study.

This study has been made possible with a grant from the Marsden fund provided by the Royal Society of New Zealand. This grant has provided funding for both the project and my living expenses.

This work has been presented at the Sixth International Conference on Advanced Materials and Nanotechnology (AMN6) in Auckland (11th-15th February 2013) and at the Second Physics of Biology Meeting in Geneva, Switzerland (26th-28th November 2013. Attendance at these conferences would not have been possible without funding from Massey University Institute of Fundamental Sciences and the New Zealand Institute of Chemistry, Manawatū Branch.

Lastly I would like to thank my partner Tina and my family who have continually supported me in my never ending quest to remain at university.

Contents

Figures and Tables

Figures

Tables

List of Nomenclature and Abbreviations

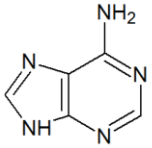
Nomenclature

The building blocks of DNA/ RNA

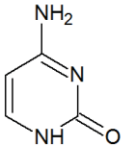
- Nucleobases

Common Bases

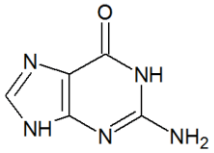
- Adenine



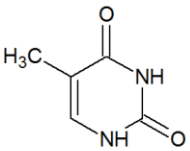
- Cytosine



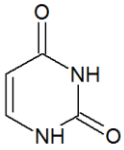
- Guanine



- Thymine

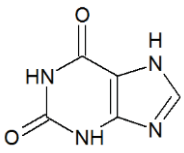


- Uracil

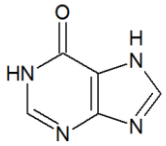


Other Bases

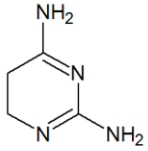
- Xanthine



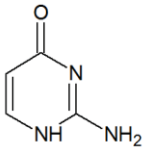
- Hypoxanthine



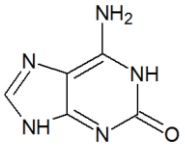
- Diaminopyrimidine



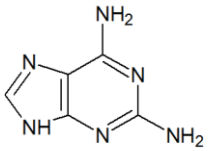
- Isocytosine



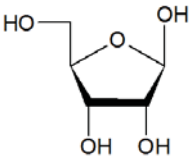
- Isoguanine



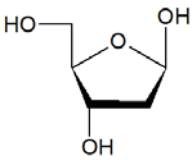
- Diaminopurine



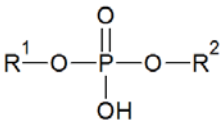
- Sugars
  - Ribose



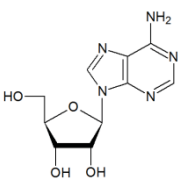
  - 2-Deoxyribose



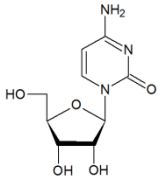
- Phosphodiester



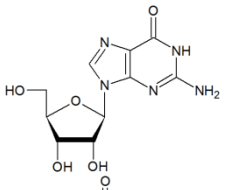
- Nucleosides
  - Adenosine



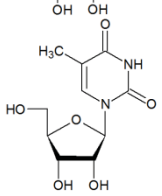
- Cytidine



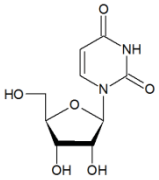
- Guanosine



- Thymidine



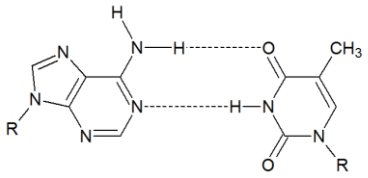
- Uridine



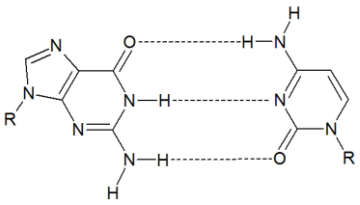
- Base pairing

  Watson-Crick

  - A-T

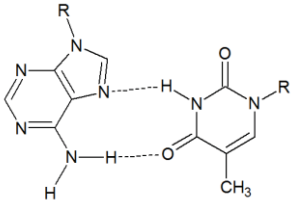


  - G-C

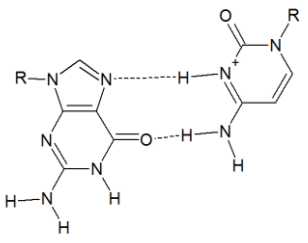


  Hoogsteen

  - A*T

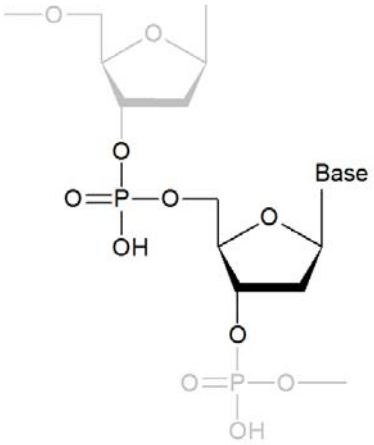


  - G*C⁺

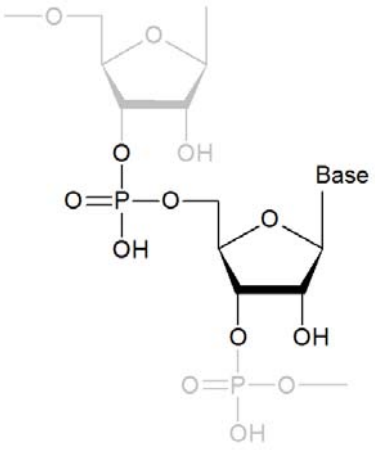


Nucleic Acids

- DNA - Deoxyribonucleic acid. Each DNA nucleotide is comprised of a nucleobase, a phosphate and a 2-deoxyribose sugar. The phosphate and sugar form the backbone of DNA. The nucleobases are attached to the anomeric carbon of the sugar unit.
    - DNA nucleotide

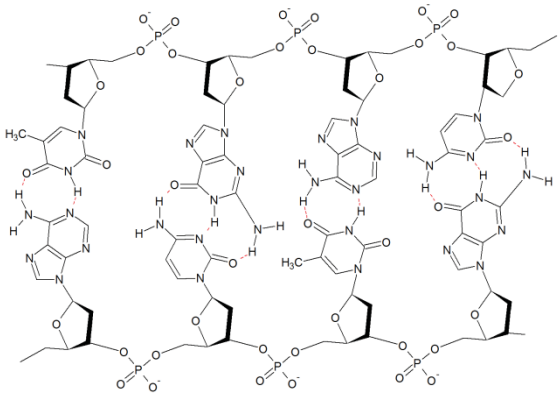


- RNA - Ribonucleic acid. Each RNA nucleotide is comprised of a nucleobase, a phosphate and a ribose sugar. The phosphate and sugar form the backbone of RNA. The nucleobases are attached to the anomeric carbon of the sugar unit.
    - RNA nucleotide



- Nucleic acid structure. A series of nucleotides joined to form a oligonucleotide chain. Watson-Crick base pairing between two

complementary oligonucleotides under standard biological conditions results in the DNA structure as show below. Three dimensionally this is a right-handed double helix with approximately 10.5 bases per turn.

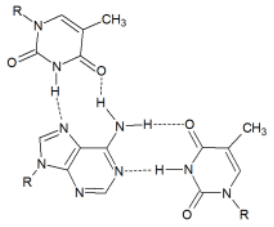


- Other DNA structures

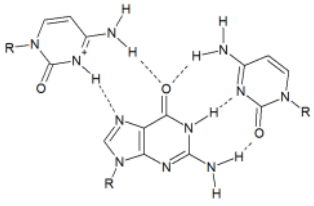
Triplexes. A triplex is formed when a C/T rich segment of a DNA sequence folds back along a A/G rich segment. This results in a triplex comprised of a pair of strands bonded by Watson-Crick pairings with a third strand bound via Hoogsteen pairings to the A/G rich strand involved in the Watson-Crick pairing.

  - Triplex pairing

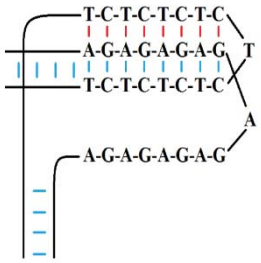
T-A*T



C-G*C+

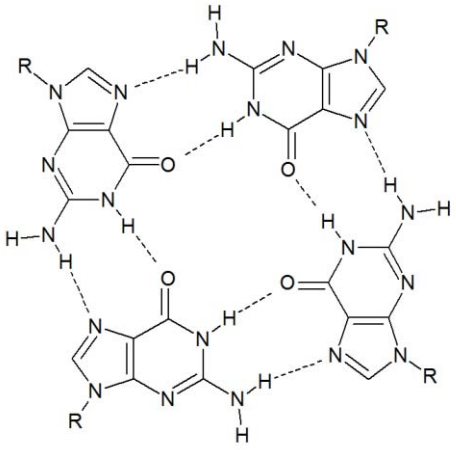


- Triplex Structure

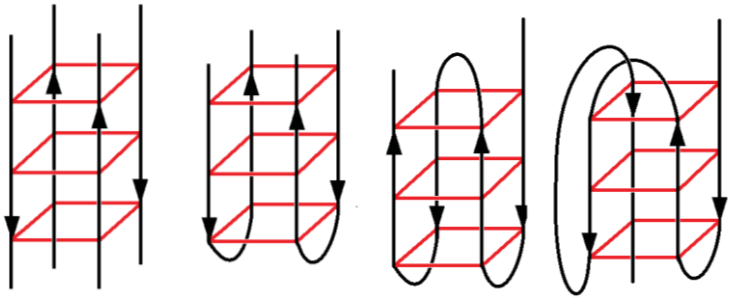


G-Quadruplexes. G-quadruplexes are a tetrad structure formed by four, G-rich sequences. Four guanine bases undergo Hoogsteen type bonding to form a G tetrad. These tetrads stack to form a quadruplex with multiple topologies.

- G-quadruplex tetrad

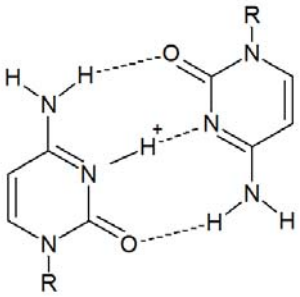


- G-quadruplex Topologies

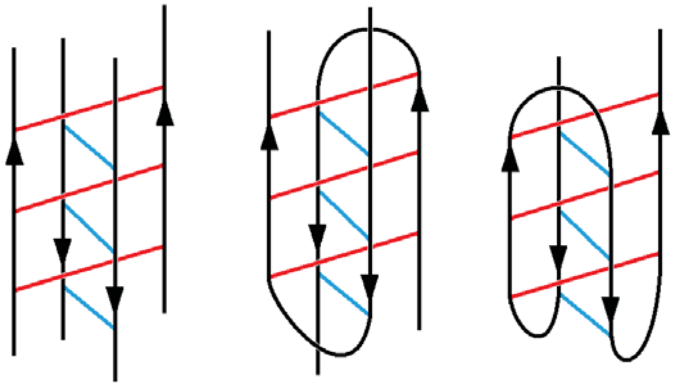


i-Motifs. i-Motifs are a tetrad structure formed by four, C-rich, DNA sequences. The structure is formed from two sets of antiparallel stranded duplexes connected by C-C^+ Hoogsteen pairings. These pairings are intercalated between the two duplexes resulting in the four stranded structures.

- i-Motif pairing



- i-Motif structures



Abbreviations

CD spectroscopy	- Circular Dichroism spectroscopy
COSY	-Correlation Spectroscopy
dNTPs	- deoxyribonucleotide Triphosphates
HiP	- High Pressure Equipment Company
NMR spectroscopy	- Nuclear Magnetic Resonance spectroscopy
NOESY	-Nuclear Overhauser Effect Spectroscopy
RF	- Radio Frequency
SBS	- Sequencing By Synthesis
Sequence 121	- DNA sequence 5′-CAT TTA TAA ATG-3′. 16.7 % G-C
Sequence 122	- DNA sequence 5′-CAT TCT AGA ATG-3′. 33.3 % G-C
Sequence 123	- DNA sequence 5′-CAA GTC GAC TTG-3′. 50 % G-C
Sequence 124	- DNA sequence 5′-CAG GTC GAC CTG-3′. 66.7 % G-C
Sequence 125	- DNA sequence 5′-CAC CCG CGG GTG-3′. 83.3 % G-C
T_m	- Melting temperature
TMSP-d	- 3-(trimethylsilyl)-2,2',3,3'-tetradeuteropropionic acid (NMR reference)

Chapter 1 - Introduction

1.1 Overview

Whilst there are many competing theories on the origin of life, a complete picture has yet to be determined which will provide the environmental, chemical and physical requirements for the evolution of stable, self-replicating, chemical systems of increasing complexity.

There is now a general agreement [1, 2] (with some exceptions [3, 4]) that, early in the development of life, RNA played the role of information storage and translation along with the role of catalysis (as ribozymes). Amino acids and peptides will have played secondary metabolic roles. By the time of the Last Universal Common Ancestor [5] to the modern domains of life, Eukarya, Archaea and Eubacteria [6], the roles of RNA had been largely usurped, with proteins playing the major role of catalysis (as enzymes) and DNA taking over the role of information storage. There is strong evidence to support this process found in the remnant RNA-mediated processes common to all current domains of life (primarily the processes within ribosomes involving ribozymes, tRNA and mRNA). These are processes which have yet to be completely replaced by proteins [7-9].

In vitro evolution experiments on the key functions required of RNA in an RNA world also support the possibility of RNA preceding DNA and proteins. Studies have shown that RNA is capable of ligase activity [10], which is essential to the generation of polymeric RNA sequences. Anabolic autocatalysis within an RNA system has also been observed [11] with incremental increases in autocatalysis with molecular size providing a way towards greater complexity. Evolution of RNA ribozymes competing in environments with limited resources has been observed *in vitro* [12]. Mutations have allowed two ribozymes to develop traits, which allow them each to exploit a different preferred resource, enabling coevolution. This shows the ability of one ribozyme to adapt to the environment without the elimination of the other ribozymes present.

These discoveries, while incredibly important, still do not answer the important question of where did life originate? This question in itself is probably unanswerable as, while some environments have an increased probability of being the crucible of life, it is impossible to know which environment was the actual location of the first life forms. There are many different possible environments, which can provide suitable conditions for the development of life, some on earth, some extra-terrestrial. And even though some of these environments may be better suited than others, self-replicating systems may still have developed in a less suitable environment. This leaves the better question to ask: what are the optimal environmental, chemical and physical requirements for the evolution of stable, self-replicating systems? That is, what is the chemical and physical environment which provides the best chance of developing life? With this in mind, the effects of various physical and chemical conditions on the stability (both physical and chemical) of RNA and DNA and their components can be probed and assessed.

1.2 Popular Theories

There are many popular theories currently in existence which describe the physical and chemical conditions which are likely to have provided the ideal environment for the evolution of self-replicating systems. This has led to numerous variables (i.e. temperature, pH, salts, salt concentration, pressure, crowding and the presence of other organic molecules to name a few) which need to be studied independently and together to determine an ideal environment.

A few of the more popular proposals cover the different possibilities of the major environmental factor of temperature with the associated environments which are or have been present on earth. These temperature environments can be categorised as: hot-start temperatures (greater than 80 °C), warm-start temperatures (between 80 °C and 4 °C) and cold-start temperature (below 4 °C). A summary of the likely environments which would provide these temperature conditions is given below.

1.2.1 Cold-Start Theories

Cold-start origin-of-life theories involve the development of life at temperatures of 4 °C or lower. These conditions would most likely be generated on a portion of the earth in an ice age event or on a global scale during a so-called "snowball" earth [13]. Even today there are many Antarctic subglacial lakes, such as Lake Vostok, which have remained isolated from the atmosphere for hundreds of thousands of years [14]. A cold-start scenario can involve both low- and high-pressure environments. These low-temperature conditions solve the enormous problem of RNA degradation, which is rapid under higher temperature conditions. However, the low-temperature environment can lack the required energy to enable many chemical reactions.

Kazakov and Vlassov has shown that the hairpin ribozyme is capable of cis and trans ligation under freezing conditions in the absence of Mg^{2+} (which is typically required for ligation in aqueous solutions at room temperature) [15-17]. At these temperatures, the backbone of the RNA chain is less susceptible to chemical instability than at higher "warm temperatures". However, the rate of ligation reaction is several orders of magnitude lower than that in a standard solution condition at biological temperatures (37 °C).

1.2.2 Warm-Start Theories

Warm-start origin-of-life theories are centred around warm mineral-rich environments such as those produced by warm alkaline springs [18] (Figure 1.1) at typical temperatures tolerable to most current organisms. These geothermal springs provide important chemicals (carbon dioxide and molecular hydrogen), in forms which are useful for the creation of bio molecules. These alkaline springs also provide micro-porous structures within calcium mineral deposits, which create cell-like environments that could separate developing systems from the hazards of the outside world [19]. The most commonly observed of these springs

are typically located on the surface of the planet and thus, they are only under atmospheric pressure. However, these springs can also be found in submarine environments (calcium mineral deposits known as "tufa" (Figure 1.2) are formed by these submarine springs), allowing for the possibility of high-pressure conditions to influence any development of life. The temperature conditions of these springs provide compensation between the drawbacks of both high and low temperatures. The temperatures are high enough to provide sufficient energy for chemical reactions but not so high as to cause rapid degradation of any potential replicating systems.

Figure 1.1: The Fly Geyser, Fly Ranch Nevada. This warm alkaline spring provides suitable conditions for the growth of red and green algae on the calcium mineral deposits [20]. Image obtained from http://exotic-place.com/wp-content/uploads/2013/12/Fly-Geyser-Nevada3.jpg.

Figure 1.2: Tufa Formations at Mono Lake, California. These formations are created by calcium deposits from submarine alkaline springs. A receding water level has now exposed this formation [21]. Image obtained from http://img.phombo.com/img1/photocombo/2711/cache/Tufa_Towers_with_Sierra_Mountains_at_Twilight_Mono_Lake_California_display.jpg.

1.2.3 Hot-Start Theories

A popular view for a hot-start origin-of-life involves deep-sea "black smokers"[22] (Figure 1.3). These submarine hydrothermal vents belch forth high-temperature gases containing a large quantity of different chemicals and heat, making them both energy- and chemical-rich. This gives a potentially ideal environment for the creation of self-replicating systems. These hydrothermal vents have been discovered at a variety of ocean depths (from 0 m to 4100 m [23]) with pressures of up to 41.3 MPa and temperatures up to 400 °C [24]. While these high-temperature conditions can provide the energy required for chemical reactions, they may also substantially increase the level of physical and chemical instability of RNA and its composite molecules. The effects of high pressure on the chemical and physical stability of RNA and its components at high temperatures are unknown.

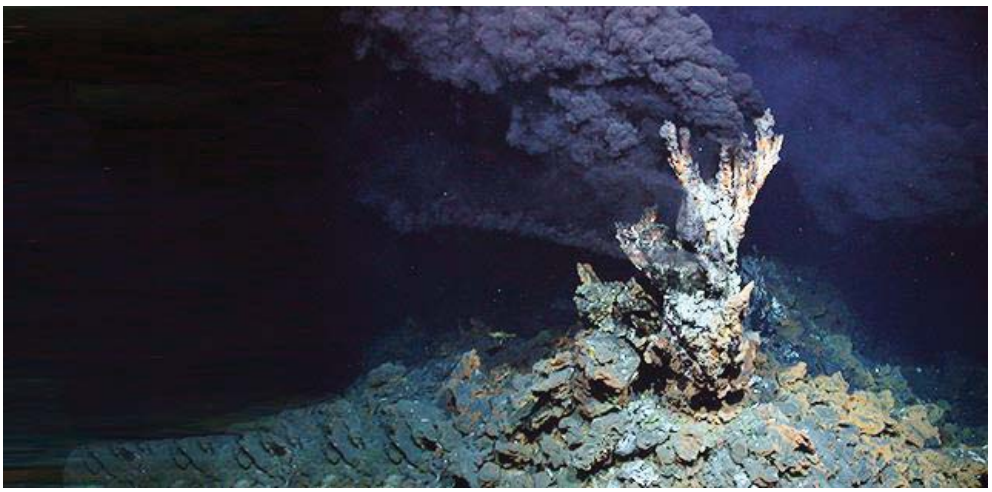


Figure 1.3: A Submarine Hydrothermal Vent. This black smoker provides energy and minerals for the surrounding life forms. Image obtained from http://www.mpi-bremen.de/Binaries/Binary5561 /black_Smoker1.jpg.

1.3 Effects of Extreme Conditions on Biological Molecules

1.3.1 High-Temperature Effects

It is common knowledge that the chemical stability of most molecules decreases with an increase in temperature. The key question to be asked is to what extent?

There are many different hyperthermophiles living at temperature extremes of greater than 80 °C [25]. This, however, is only possible as a result of *in vivo* evolution producing RNA sequences with enhanced thermal stability [26]. This, along with enzymes that repair instances of hydrolysis within the DNA/RNA chains, enables the possibility of life at temperature extremes. In the absence of these mechanisms, DNA and RNA are inherently unstable under high-temperature conditions.

Studies by Levy and Miller examined the temperature dependence of the chemical stability of individual RNA and DNA nucleobases [27]. It was discovered that at temperatures of 350 °C, the half-lives for hydrolysis of most bases, in neat form, were between 2 and 15 sec. They observed that at 100 °C in aqueous solutions most bases had half-lives between 0.8 and 56 years, while cytosine has a half-life of only 19 days. These short half-lives indicate that most building blocks of RNA/DNA exist only for a short time in high-temperature environments with cytosine being the least stable.

A study by Frederrico *et al.* observed that the average rate of hydrolysis of two cytosine residues within a single-stranded piece of DNA was approximately the same as for the individual base [28]. They also observed that the average rate of hydrolysis of these two cytosine residues within a double-stranded piece of DNA was approximately 140 times slower than the rate in the single-stranded DNA.

In addition to this chemical instability of the individual nucleobases, the RNA backbone itself is also chemically unstable at elevated temperatures. At temperatures approaching 100 °C, the backbone itself undergoes facile hydrolysis, breaking the phosphodiester bonds [29]. With this low chemical stability at elevated

temperatures, the ability of DNA/RNA to translate a genetic code is significantly reduced.

Along with this decrease in the chemical stability at high temperatures, the physical stability of DNA/RNA is also reduced. It has been shown [30] that the 2D and 3D structures of RNA molecules, in the absence of proteins, are unstable at elevated temperatures resulting in the loss of catalytic function. Additionally, it is well known that both DNA and RNA strands are unable to bind to other DNA/RNA fragments at elevated temperatures, making the reliable copying of genetic code impossible under high-temperature conditions.

These results clearly show that high temperatures have a large negative effect on both the chemical and physical stability of RNA and its constituents, particularly as temperatures approach 100 °C.

1.3.2 High-Pressure Effects

The chemical and physical stability of oligonucleotides (both DNA and RNA in their various forms) at extremes of pressure remain largely unknown [31-33]. A study by Wilton *et al.* [34] has shown that, at high pressures, the volume of a small DNA hairpin changes little; rather that the shape of the DNA changes to reduce the volume of the hydration layer (this can be seen as a widening of the minor grove). While this is insightful, it does not address whether there is any significant change in the DNA's physical susceptibility to temperature.

By contrast, the effect of high pressures on the physical stability of proteins [35, 36] and even whole organisms [37] has generated enormous interest. Early studies by Nash and Jonas [38] showed that high pressure could be used for the cold denaturation of egg-white lysozyme. Later experiments have showed that proteins undergo a variety of different behaviours, from an increase in structural stability [39], to complex local denaturation effects which vary considerably with pressure and structure [40].

However, for both DNA/RNA and proteins, the effects of high temperatures and high pressures have only been studied along the axes of the pressure-temperature landscape. This has result in a large gap in the current knowledge of the effects of combined high-temperature/high-pressure environments on the chemical and physical stability of DNA/RNA.

Irrespective of hot or cold start, at atmospheric pressures or under high-pressure conditions, fundamental physical and chemical data at extremes of both pressure and temperature are still urgently needed on key molecules (such as nucleobases and oligonucleotides), in order to offer evidence for or against various hypotheses for the origin-of-life.

1.4 Aims of this Thesis

The overall aim of this thesis is to provide the first chemical and physical studies examining the role of both pressure and temperature on chemical and physical stability of DNA and its constituents using predominantly NMR spectroscopy. This will fill in some of the large gaps in the understanding of the effects of high pressure on the chemical and physical stability of DNA, particularly the effects of both high temperature and high pressure. The outcomes of this research will be used to improve the current understanding of the potential role of high-pressure on the origin of self-replicating systems. It will also address some fundamental questions: Will high pressures have any effect on the chemical or physical stability of DNA? If so, will this be a positive or negative effect? And is this change in stability capable of sufficiently offsetting the destabilising effects, both chemical and physical, of high temperature on DNA and its constituent nucleobases?

Examination of the effects of high pressure on the chemical and physical stability of nucleic acids will be performed using DNA. The difficulty in removing RNases from equipment makes RNA extremely difficult to work with and will most likely lead to the loss of expensive RNA samples. DNA will be a suitable, more stable, alternative. The similarities of DNA to RNA (particularly the physical structure of

small duplexes) will allow the results obtained for DNA to be considered as acceptable analogues for RNA until more results are obtained.

High-pressure NMR spectroscopy will allow for the observation of the chemical composition of solutions as well as being able to observe physical behaviours as changes in chemical shift and intensity of NMR signals. This technique also has a great advantage in being able to observe the behaviours of each individual nucleotide in a DNA sequence rather than only observing the average behaviour as with other high-pressure spectroscopic techniques (such as UV-visible spectroscopy and circular dichroism spectroscopy).

This aim can be broken down into four major parts, which will each be independently studied. Detailed introductions are given within each of the research chapters, providing further information relevant to the enclosed studies.

1.4.1 Development of a High-Pressure System

To study DNA under high-pressure conditions, a specialised high-pressure system will be developed. This system will have to allow for the safe-handling and operation of a commercial high-pressure NMR cell as well as providing a mounting point for a high-pressure pump which will pressurise the whole system. A removable high-pressure reaction vessel will also be added to this system to extend its capabilities beyond the NMR spectrometer.

1.4.2 Chemical Stability of Individual DNA Bases

High-pressure NMR spectroscopy will be used to examine the effects of high pressure on the rate of hydrolysis of cytosine and cytidine (the least stable of the nucleobases and therefore a limiting factor) at 100°C. This will involve examining the rate of change of cytosine/cytidine concentrations as they decay, to uracil/uridine respectively (Figure 1.4). This will be repeated under multiple

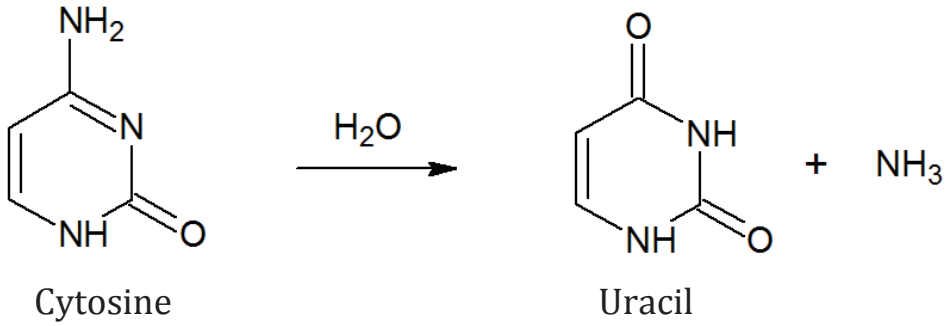


Figure 1.4: Equation for the Hydrolysis of Cytosine to Uracil.

pressure conditions to establish the reaction volume of cytosine/cytidine hydrolysis. A brief assessment of the rate of cytosine hydrolysis as a function of pH at both high and low pressures will also be performed.

1.4.3 Physical Stability of DNA Structures

High-pressure NMR spectroscopy will be used to examine the effects of high-pressure on the melting points of a self-complementary DNA hexamer and five self-complementary DNA dodecamers. This will first involve the use of NOESY and COSY ^{1}H NMR techniques to assign the peaks of the ^{1}H NMR spectra for each of the DNA sequences. This will be followed by melting experiments which can be used to determine melting curves, and therefore melting points, for, in principle, each residue of each sequence under multiple pressure conditions.

The effect of high pressure on the physical stability of a DNA i-motif will also be examined. The melting point of the i-motif sequence will be determined using high-pressure NMR spectroscopy under different pressure and pH conditions.

These studies will develop an understanding of how the physical stability of the DNA sequences is affected by high-pressure.

1.4.4 Chemical Stability of DNA

Next-generation sequencing technology will be used to determine the effect of high pressure and high temperature on the chemical stability of cytosine within both single-stranded and double-stranded bacteriophage DNA. Samples will be incubated under a variety of temperature and pressure conditions after which they will be sequenced to determine any cytosine hydrolysis (which will be seen as a cytosine to thymine transition) within the DNA sequence. With this ability to examine each cytosine in the sequence it will be possible to study the effect of high pressures on the chemical stability of each cytosine residue in the DNA sequence and to determine whether sequence position results in any differences in the chemical stability and the susceptibility of the chemical stability to pressure.

1.5 Summary of Thesis Structure

Chapter 2 will provide an overview of the high-pressure system. This will include information on high-pressure NMR cells, as well as information on the construction and operation of the safe-handling apparatus, the pump system and the high-pressure reaction vessel. A brief discussion of the effects of high pressure on buffer pH is also included.

Chapter 3 includes the results of studies on the effects of high pressure on the chemical stability of cytosine and cytidine. This will involve hydrolysis measurements for both cytosine and cytidine at multiple pressures between 0.1 MPa and 200 MPa. These hydrolysis measurements will be used for rate calculations and determination of reaction volumes of these bases.

Chapter 4 includes the results of studies on the effects of high pressure on the physical stability of a self-complementary DNA hexamer. This will involve determination of the melting point of the DNA under multiple pressure conditions between 0.1 MPa and 240 MPa. These melting points will be determined from melting curves derived from the changing chemical shifts for each nucleotide

within the sequence. The methodology for the assignment of the ^{1}H NMR spectra of the DNA is also included. This work will be extended to cover five self-complementary DNA dodecamers of varying G-C content in **Chapter 5**.

Chapter 6 includes the results of studies on the effects of high pressure on the physical stability of a DNA i-motif sequence. This will involve determination of the melting point of the i-motif under multiple pressure conditions between 0.1 MPa and 200 MPa at pH levels between pH 4.1 and pH 7. These melting points will be determined from the disappearance of the signal corresponding to the H^{+} involved in the $C\text{-}C^{+}$ bonding of the i-motif structure.

Chapter 7 describes the methodology involved in the preparation and incubation of samples which will be used to determine the effects of high pressure on the chemical stability of cytosine within a DNA sequence. This will include incubation techniques, PCR methods and results indicating sample quality. Also included will be a description of the next-generation sequencing technique which will be used to sequence the incubated DNA to identify hydrolysis events. Sequencing of samples has yet to be performed and, as such, has prevented a full presentation of these results in this thesis.

Chapter 2 - The High-Pressure System

2.1 High-Pressure NMR Cell

High-pressure NMR studies have had a long and detailed history, but one which has traditionally been limited to a select few institutions with custom-made high-pressure NMR cells. These cells have been fabricated by these institutions and have undergone years of specialist development. While the performance of high-pressure NMR cells has made great strides over the course of their history, there have always been numerous limitations brought upon by the fact that the use of metal tubes within an NMR spectrometer is simply not an option. Development of the high-pressure cells can be categorised into one of four major techniques [41]: fused silica capillaries, metallic pressure vessels, diamond anvil cells or large-volume tubes.

2.1.1 Fused Silica Capillaries

Fused silica capillaries (Figure 2.1) are a relatively simple approach to high-pressure NMR [34, 40, 42]. This method uses long lengths of fused-silica capillary tubing, usually those used by the chromatography community, to create a high-pressure NMR cell. The tubing is folded back and forth repeatedly, using a flame, to increase the volume of sample residing in the NMR spectrometer's radio frequency (RF) coil. The folded capillary is coated in a protective layer and usually placed in a standard NMR tube for protection. Even with defects from the manufacturing process, cells made with this method can withstand pressures of up to 100 MPa without additional special treatment. Yamada *et al*. has performed extensive research into glass capillary NMR cells [43]. They describe extremely thorough methods of etching and cleaning the glass to avoid stress risers and have produced cells with thick-wall capillaries withstanding pressures up to 200 MPa. These cells, as impressive as they are, are still limited by their low sample volume, requiring

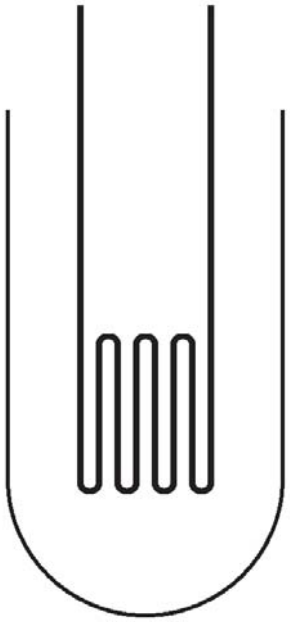


Figure 2.1: Fused Silica Capillary Cell. The capillary is bent back and forth several times to increase the sample volume within the RF field. it is then placed within a standard NMR tube for protection. Image adapted from Conradi [41].

smaller more concentrated samples. This can lead to solubility issues, aggregation and reduced signal-to -noise ratios.

2.1.2 Metallic Pressure Vessels

Metallic pressure vessels are composed of a strong, nonmagnetic casing which contains the hydraulic fluid [44] (Figure 2.2) with the RF coils located inside the pressure vessel since the RF field is unable to penetrate through the walls of the metallic vessel. The sample vessel is then placed within the RF coils and pressurised directly or indirectly by the surrounding hydraulic fluid. There has been a large number of these high-pressure cells developed reaching pressures of 700 MPa. These cells are also capable of much larger volumes than capillary cells. The installation of a pressurised vessel containing the RF coils into an NMR spectrometer probe is a non-trivial task, requiring expert knowledge of NMR spectrometer construction. In addition to this, any failure of the pressure vessel will cause severe damage to the NMR spectrometer. A further drawback for this

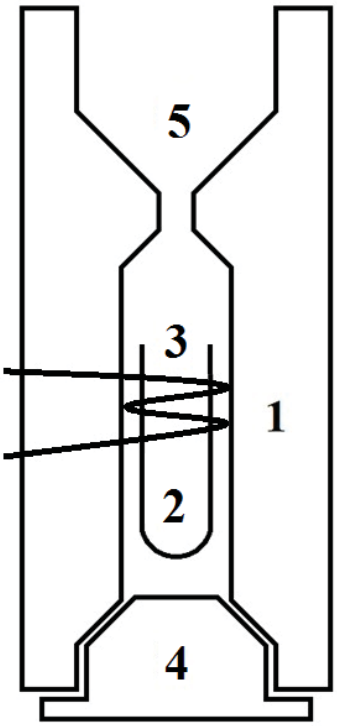


Figure 2.2: Metallic Pressure Vessel. This cell is constructed from a metal pressure vessel (1) containing the sample tube (2) which is placed inside the RF coil (3). The vessel is sealed from the bottom with a plug (4) and pressure is provided from an external pressure source connected to the top of the cell (5). Image adapted from Conradi [41].

type of cell is the spectral interference from the hydraulic fluid which surrounds the RF coil. While this can be avoided with the use of liquids with no hydrogen spins, these fluids are usually expensive or hazardous.

2.1.3 Diamond Anvil Cells

Diamond anvil cells [41] for use in high-pressure NMR measurements use direct contact with the anvils to generate pressure (Figure 2.3). A gasket material is used to prevent the sample from escaping radially outward. This cell is mounted on top of a NMR spectrometer probe and surrounded by an RF coil. A typical problem with these cells is shielding of the RF field when using metallic gaskets. Special coils can be made which create field lines, which dip into and out of the sample

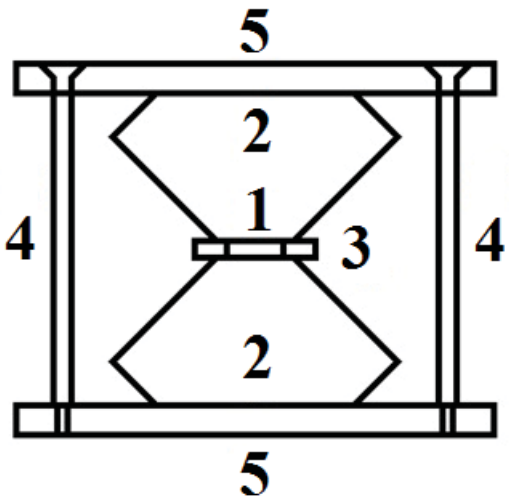


Figure 2.3: Diamond Anvil Cell. The sample is placed in the sample region (1) between the two diamonds (2). A metallic gasket (3) holds the sample in place. The RF coil is situated around this middle region of the cell. Pressure is applied by tightening the screws (4) which connect the backing plates (5). Image adapted from Conradi [41].

region, avoiding any closed conducting paths. Diamond anvil NMR cells can generate pressures of the order of 10,000 MPa, far beyond the capabilities of other high-pressure NMR techniques. This cell is also ideal for use in solid-state NMR spectroscopy as it can apply direct mechanical pressure to the sample. However, the very small sample volumes severely limit the potential of this cell for liquid-phase NMR spectroscopy.

2.1.4 Large-Volume Tubes

Large-volume high-pressure NMR tubes are an attempt to construct a high-pressure vessel in the form of a traditional NMR tube attached to a manifold (Figure 2.4). Pressure is applied to the sample either via direct contact between the hydraulic fluid and an immiscible sample or indirectly via a Teflon piston or diaphragm. These tubes are traditionally made from single-crystal sapphires and can have internal cross-sectional areas of up to half that of an equivalent standard 5 mm NMR tube while maintaining the same external dimensions. This setup allows for large volume samples to be studied with the use of conventional NMR

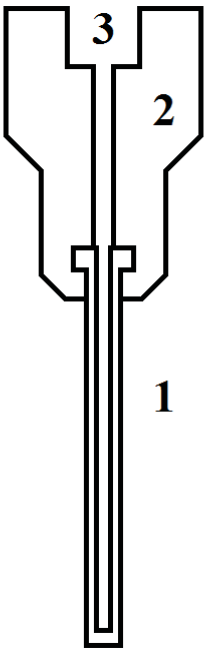


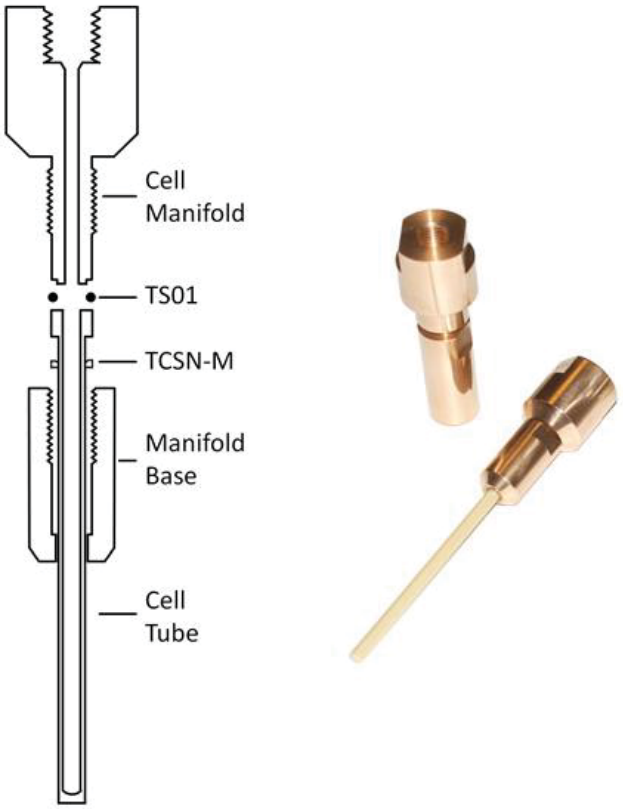
Figure 2.4: High-Pressure NMR Tube. The sample is placed in the tube (1) which is connected to a manifold (2). The cell is pressurised by a pump connected to the high-pressure connection point (3). Image adapted from Daedalus Innovations [45].

spectrometer probes and even cryoprobes. However, in the past, there have been limitations on the pressures that can be achieved with this technique. The sapphire tubes originally developed by Roe [46] are capable of up to 100 MPa with only 15 MPa as a safe operating pressure. Recently there have been great advances made in the pressure capabilities of these cells with the use of ceramic [47] tubes which are capable of safe operating pressures of up to 250 MPa.

2.1.5 The Daedalus Innovations Cell

With research conditions requiring large volumes and without the resources, expertise or time to construct a metallic pressure vessel, the large volume tube technique was chosen for the construction of the high-pressure cell. The chosen cell is a ceramic tube system developed by Daedalus Innovations [45]. The high-pressure cell is constructed from a two-part aluminium manifold (later replaced with a titanium equivalent) which secures a zirconia high-pressure NMR tube (Figure 2.5). The high-pressure tube is 92 mm long and has an outer diameter of 5

mm and an internal diameter of 3 mm (standard NMR tubes have an internal diameter of approximately 4.2 mm). This gives an internal cross sectional area, and therefore sample volume, of approximately half of that of a standard NMR tube. The zirconia construction has the added benefit of a tube that is considerably more robust than tubes constructed from most other materials. It is reported from Daedalus Innovations that the tube is capable of being dropped from a lab bench without loss of its high-pressure capabilities. The cell setup is capable of withstanding pressures of up to 250 MPa with an operational temperature range of 5˚C – 100˚C. This temperature range is a result of the physical capabilities of the nitrile seal between the tube and the manifold. New seals are now available that will extend the operational temperature range to -15˚C – 125˚C. The high-pressure cell is connected to an external pressure generator by means of a high-pressure tether attached to the top of the cell. This connection gives the cell an on-line capability allowing the pressure within the cell to be changed and monitored without the need to remove the cell from the NMR spectrometer. The sample within the cell is separated from the hydraulic fluid using an immiscible liquid interface, typically an aqueous sample with a hydrophobic hydraulic fluid. This setup allows the study of large volume samples under high-pressure/high-temperature conditions without any modification to existing NMR spectrometer equipment and with the ability to monitor and alter pressure easily.



(a) (b)

Figure 2.5: The Daedalus Innovations Cell. (a) A schematic of the Daedalus cell. The ceramic tube is secured in the cell manifold by the manifold base. A nitrile O-ring (TS01) provides the seal while the metallic tube seat (TCSN-M) ensures correct seating of the tube. Pressure is provided via a high-pressure connection at the top of the cell. Image adapted from Daedalus Innovations. (b) The cell manifold and the completed cell. Image obtained from http://www.daedalusinnovations.com /apparatus/high-pressure.html.

2.2 Safe-Handling Apparatus for the High-Pressure NMR Cell

Due to the very high pressures involved, any shocks to the high-pressure tube during operation while the cell is under pressure could result in a catastrophic failure of the cell. Cell failure will endanger the operator along with any other persons in the general vicinity. The likelihood of this occurring is increased by the rigidity of the high-pressure tether which makes the handling the cell a difficult task. For the sake of safety of both the operator and the cell, it was decided that a safe-handling apparatus be constructed to which the cell can be attached to reduce all likelihood of cell failure.

This safe-handling apparatus has to conform to three major requirements:

- First, it must allow for a shaft to be secured to the top of the NMR spectrometer. The high-pressure cell is attached to the end of this shaft enabling it to be inserted and removed from the NMR spectrometer in a controlled reproducible fashion.
- Second, the apparatus must be able to move the cell to an appropriate location for the assembly and disassembly of the cell. This location must be a sufficient distance from the NMR to allow for the use of ferromagnetic tools.
- Finally, the apparatus must be constructed from non-ferromagnetic materials due to the strong magnetic fields generated by the NMR spectrometer.
- Optionally, the safe-handling apparatus can provide a mounting platform for the high-pressure pump system required to pressurise the cell.

The result of these requirements is a wall-mounted arm, which is capable of positioning a shaft, to which the high-pressure cell is attached, over the NMR spectrometer. This arm is capable of positioning the cell a safe distance from the NMR spectrometer where it can be assembled and disassembled. A section of the arms frame was designed to allow for the mounting of the high-pressure pump system. The apparatus is constructed from aluminium with stainless steel fittings. The following is a list of the key features of this design with references to their location on the schematic shown in Figure 2.6.

1. High-pressure cell. This is the position of the high-pressure cell at the base of the shaft.
2. Aluminium shaft. The high-pressure cell is attached to the shaft via a small PEAK plastic fitting. The length is such that, when lowered into the NMR spectrometer, the top of the shaft is sits on top of the shaft sleeve reducing the weight of the shaft on the NMR probe. The high-pressure tether passes through the shaft and exits out the top.

3. Shaft sleeve. This sleeve allows the shaft to be positioned at the desired height and secured with the use of the locking nut located at the top of the sleeve. A fine adjustment collar is located on the outside of this sleeve so that its position within the alignment bracket can be changed.

4. Alignment bracket. This bracket has been developed so that the shaft can be easily and reproducibly positioned in an exact location over the top of the spin housing at the top of the NMR spectrometer. Three locking screws are positioned around this bracket so it can be secured to the spin housing in order to limit any movement of the high-pressure cell. When the cell is drawn up into the bracket prior to insertion into the NMR spectrometer, the cell is only visible from directly below the bracket. This shields the operator from any possible cell failure and makes it virtually impossible to strike the cell against objects.

5. Main arm. The main arm provides the mounting point for the high-pressure pump system. It has a full range of motion from the NMR spectrometer to flat up against the wall of the NMR lab.

6. Secondary arm. This extension to the main arm allows the alignment bracket to be manoeuvred around the nitrogen ports on top of the NMR spectrometer. It also contains two adjustment features which ensure that the shaft is positioned vertically.

7. High-pressure pump system mounting position.

8. Wall-mounting bracket

9. Locking arm. This arm can be attached to the main arm to lock it in position so the high-pressure pump can be operated without movement of the apparatus.

10. High-pressure tether. This tether transfers the pressure from the high-pressure pump system to the high-pressure cell. The tether has had a coil added to allow for the movement of the shaft up and down.

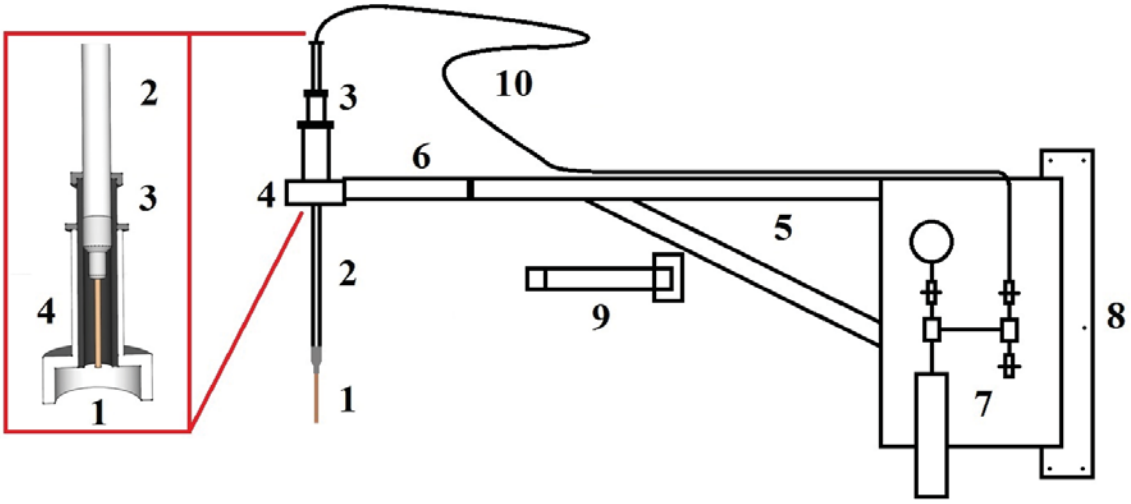


Figure 2.6: Safe-Handling Apparatus for the High-Pressure NMR Cell. This schematic outlines the important features of the safe-handling apparatus. Further information on these features is given in the text. (1) The high-pressure cell. (2) Aluminium shaft. (3) Shaft sleeve. (4) Alignment bracket. (5) Main arm. (6) Secondary arm. (7) High-pressure pump. (8) Wall-mounting bracket. (9) Locking arm. (10) High-pressure tether.

This apparatus was designed in consultation with and was custom built and installed by the Massey University Institute of Fundamental Sciences (IFS) Engineering Workshop. Only minor alterations from the original design have been necessary.

2.3 High-Pressure Pump System

2.3.1 Initial Development

To achieve the high pressures required for the operation of the high-pressure cell, a high-pressure pump system had to be developed. The proposed pump system is shown below in Figure 2.7. This system was assembled with components manufactured and provided by the High Pressure Equipment Company (HiP [48]). The entire system is rated to 60,000 psi (413.7 MPa) with the exception of the pressure gauge which is rated to 75,000 psi (517.1 MPa).

The pump chosen to provide pressure for the system is the HiP high-pressure generator, extra capacity model. This is a manually operated piston screw pump

and is rated for 60,000 psi (413.7 MPa). All wetted parts are made from 316 stainless steel and 17-4PH stainless steel. The pump has a stroke length of 146.05 mm with one revolution of the handle giving 1.81 mm of travel. A total of 10 mL of hydraulic fluid is pumped per stroke.

The pressure gauge used is a HiP Bourdon gauge rated for 75,000 psi (517.1 MPa). This gauge has a solid front separating the dial face from the pressure tube in order to provide maximum safety along with a safety blowout back and a removable polycarbonate plastic front window. The Bourdon tube and socket are made from 316 stainless steel. The accuracy of the gauge is ± 2.5 MPa.

All other components (tubing, valves, tees, adapters, glands and collars) are made from 316 stainless steel. Connections between tubing and components are with a gland nut and collar arrangement as shown in Figure 2.8. This connection has become an industry standard for use at elevated pressures and temperatures in both liquid and gas applications and may be disassembled and retightened indefinitely. Technical specifications for the high-pressure pump, gauge and other high pressure components are available in the digital appendices.

This pump system is designed to allow for easy setup and operation. The system can be initially primed with the use of a vacuum pump. This is done by first closing off all the outlet valves followed by the removal of all the air in the system using the vacuum pump. The resulting vacuum is then used to draw hydraulic fluid from the reservoir into the system, eliminating any air pockets. The inlet valve can then be closed and once the sample has been inserted in the tube and the cell assembled, pressure can then be applied to the system with the use of the hand pump. Further information on the operation of the high-pressure pump system is given in the High-Pressure Pump Manual in Appendix 6. The fabrication of mounting brackets and the mounting platform was done by the Massey University IFS Engineering Workshop.

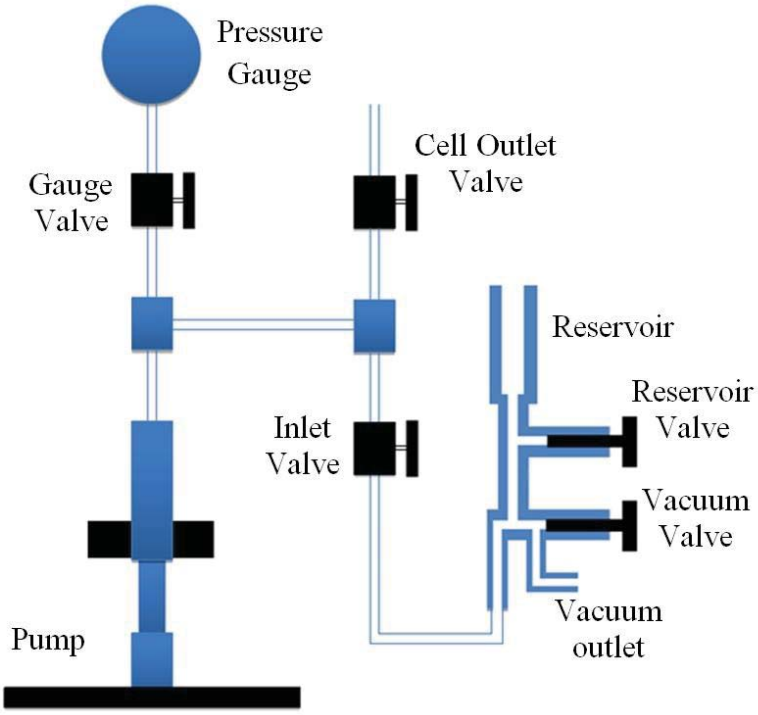


Figure 2.7: Schematic of High-Pressure Pump. The high-pressure NMR cell is connected to the cell outlet valve via the high-pressure tether connected at the cell outlet.

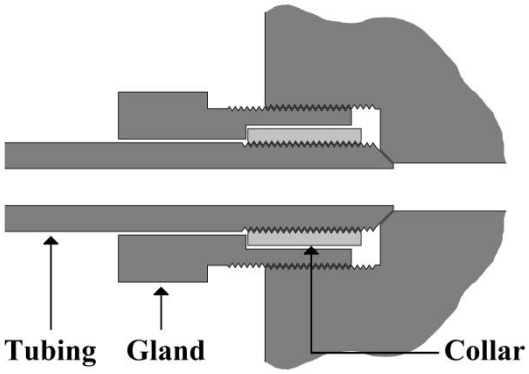


Figure 2.8: High-Pressure Connection. A collar is screwed onto the end of the high-pressure tubing. A gland nut is then threaded over the collar and screwed into the component. This forces the coned end of the tubing in the orifice of the component creating a seal. This connection may be disassembled and reassembled indefinitely. The threads of the gland are right-hand while the threads of the collar and tubing are left-hand to prevent rotation of the collar during assembly. Image adapted from HiP [48].

2.3.2 Upgrades to High-Pressure System

The later addition of a high-pressure reaction vessel meant that the capabilities of the high-pressure pump system had to be extended. This was in the form of the

replacement of the "T" junction below the pressure gauge with a four-way cross junction to allow for the addition of an external outlet valve (Figure 2.9). This valve was mounted to the back of the safe-handling apparatus. A high-pressure tether which connects the reaction vessel to the pump system was connected to the valve. Due to the weak ferromagnetic nature of the reaction vessel, a reaction vessel mounting platform has been constructed to secure the high-pressure reaction vessel to the floor of the lab. This platform prevents any potential movement of the high-pressure reaction vessel but still allows easy operation. It also provides support for the high-pressure tether. The additional high-pressure components were manufactured and provided by HiP and are made from 316 stainless steel. The bending of the high-pressure tubing and the fabrication of brackets and the reaction vessel platform were done by the Massey University IFS Engineering Workshop.

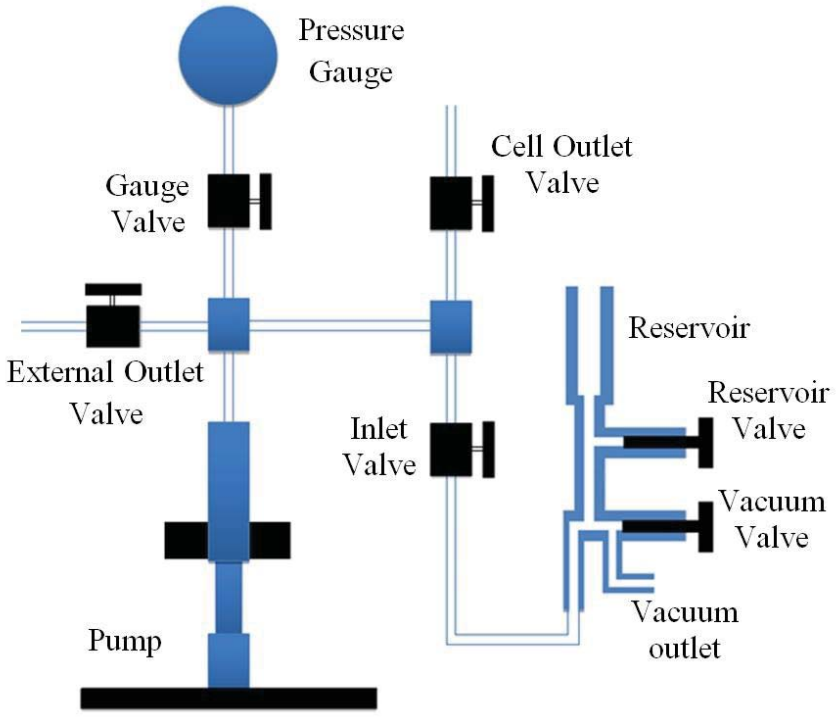


Figure 2.9: Schematic of High-Pressure Pump Upgrade. The external outlet valve connects to the high pressure reaction vessel via a high-pressure tether.

2.3.3 Safety

With the very high pressures involved in the use of this system, safety is of great concern. The high-pressure NMR cell was provided with a clear Perspex box for the containment of the cell during pressurisation. A face shield has been sourced for protection during use of the pump system. Operational guide lines have also been written to outline correct procedures for the operation on the high-pressure NMR cell and the pump system. It can also be noted that this system is relatively safe because of the use of a liquid hydraulic fluid. This results in a very small compression volume (approximately a 5 mL maximum for the high-pressure cell and a 15 mL maximum for the high-pressure reaction vessel below) as compared to systems compressing gas. Any failures within the system will lead to a rapid decrease in the system pressure with only a small volume of released hydraulic fluid. With these considerations, the potential for harm to the operator has been reduced to a near-insignificant risk.

2.3.4 Choice of Hydraulic Fluid

With the need to use aqueous samples it is clear that a non-polar hydraulic fluid would be required to provide the needed immiscible liquid interface within the high-pressure NMR cell. This hydraulic fluid would need to be stable at 100 °C, nontoxic, inexpensive, have a low viscosity and be less dense than water. Light paraffin oil was chosen to be the hydraulic fluid as it satisfied all these conditions. Tests were made to see if any unwanted NMR signals were generated by the oil while in contact with the sample. No signals were seen.

2.4 High-Pressure Reaction Vessel

In order to conduct large-scale experiments that do not require the use of NMR spectroscopy, a high-pressure reaction vessel was required. The chosen reaction

vessel is the HiP series “R” O-ring closure reactor (Figure 2.10). The reactor is constructed from H1150 stainless steel and is rated to 60,000 psi (413.7 MPa). The internal diameter of the reactor is 25.4 mm with a depth of 152.4 mm giving a total internal volume of 77.2 mL. Sealing is accomplished by a highly reliable combination of O-ring and separate metal back-up ring. The wedge-shaped back-up ring is designed to expand and contract as pressure increases or decreases. Consequently, the O-ring is continuously confined with no clearance for extrusion.

To enable a constant temperature within the high-pressure reaction vessel, it would need to be placed in a water bath during the course of its use. It was required that the reaction vessel could be isolated from the pump system, so as to allow for its detachment from the connecting high-pressure tether. To enable this, a frame was fabricated to hold the reaction vessel as well as providing a mounting point for the required isolation valve. This isolation valve is connected via a custom bent section of tubing to the pressure connection on the underside of the reaction vessel. The connection on the reaction vessel cap was sealed with a standard high-pressure plug. A small custom-made steel insert was inserted in to the cavity on the underside of the reaction vessel cap. The purpose of this insert is to prevent oil draining out of the cavity in the reaction vessel cap but to still allow oil to pass through the insert to prevent any pressure build up. This is achieved by a small hole through the centre of the insert which is small enough to stop oil flowing out under the force of gravity.

A simple capillary holder was also fabricated to hold up to 18 capillaries within the high-pressure reaction vessel. This capillary holder has the additional purpose of occupying as much space within the reaction vessel as is safely possible. This reduces the volume of hydraulic fluid in the reaction vessel, reducing the amount that would need to be compressed during pressurisation.

The bending of the high-pressure tubing and the fabrication of the reaction vessel frame, insert and the capillary holder was done by the Massey University IFS Engineering Workshop.

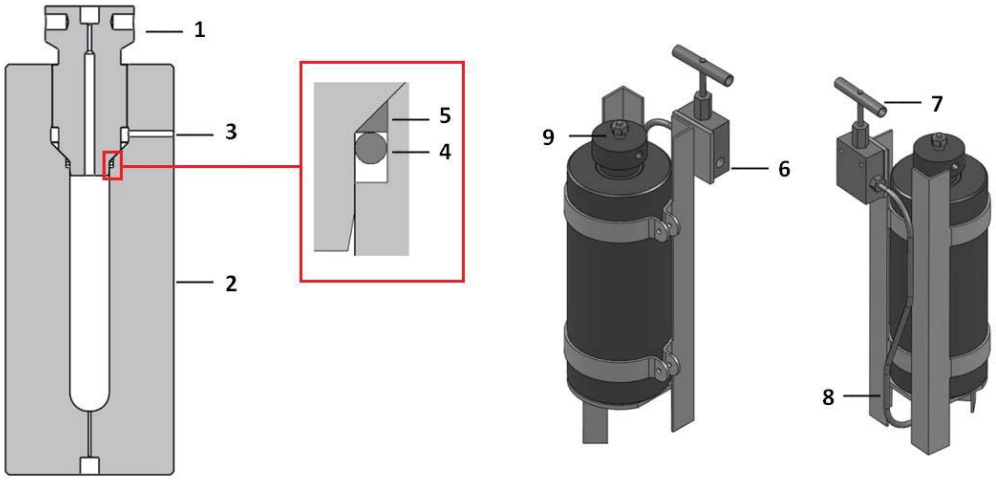


Figure 2.10: High-Pressure Reaction Vessel. (a) A cross section of the reaction vessel showing the cap (1), body (2), vent hole (3), O-ring (4) and backup ring (5). (b) The reaction vessel in its frame. This shows the high-pressure tether connection (6), the isolation valve (7) tubing to connect the valve to the underside of the reaction vessel (8) and the high-pressure plug (9).

2.5 Buffer Correction

The dependence of temperature on the pH of a buffer solution is well known and can be easily determined. However, the pressure dependence of pH is less well understood but, nevertheless, will play an important role in the course of this research. Bruins *et al.* addressed this problem and discussed the result of the pH changes on past research [49]. The pH value of buffer solutions is altered by pressure and like all equilibrium reactions this change is determined by Le Chatelier's principle. Equation (1) describes the thermodynamic relationship of the ionisation constant K and pressure P:

$$\left(\frac{\partial \ln K}{\partial P}\right)_T = -\frac{\Delta V}{RT} \tag{1}$$

T is the absolute temperature, R is the ideal gas constant and ΔV is the reaction volume of the system. The reaction volume of the system itself is pressure

dependent. It was shown by El'yanov that the relationship between the ionisation constant at atmospheric pressure, K_0, and at pressure, K_p, can be expressed as [50]:

$$\ln\left(\frac{K_p}{K_0}\right) = -\frac{P\Delta V^0}{RT(1+bP)} \quad (2)$$

where b is a universal constant 9.2 x 10^{-4} MPa^{-1} and ΔV^0 is the reaction volume under atmospheric conditions. With the use of equation (2), it can be shown that there are significant differences in the pH of a buffer solution between atmospheric pressure and 250 MPa. In the case of phosphate buffer at pH 7, the pH of the buffer solution will change by 0.1 pH units at as little as 25 MPa. These changes are significant and need to be accounted for during high-pressure experiments especially those experiments that are sensitive to pH changes. Equation (2) along with a list of known values for reaction volumes (Table 2.1) makes it possible to determine the pH change of a buffer solution under high-pressure conditions. This will allow a corrected atmospheric pH value (pH_0) to be determined which will compensate for the pH shift to produce the correct pH under pressure. Table 2.2 displays the pH shifts of a pH 7 phosphate buffer along with the pH_0 value at commonly used pressure levels.

Buffer	**pK_a**	**ΔV^0 /cm^3 mol^{-1}**
citrate	4.55	-12.3 ± 0.4
MES	6.10	3.6 ± 0.1
Cacodylate	6.17	-13.2 ± 0.2
Phosphate	6.70	-24.0 ± 0.4
HEPES	7.50	4.8 ± 0.1

Table 2.1: Buffer Reaction Volumes (ΔV^0). This table shows the pK_a and reaction volumes [51] of buffers relevant to this research.

Pressure /MPa	pH shift	pH_0
0.1	0	7.00
50	-0.20	7.20
100	-0.39	7.39
150	-0.56	7.56
200	-0.71	7.71
250	-0.86	7.86
413	-1.27	8.27

Table 2.2: Buffer pH Corrections. Corrected atmospheric pH values (pH_0) for pH 7.0 phosphate buffer for commonly used pressure levels at 298.2 K.

2.6 Problems and Failures

This high-pressure NMR cell has operated without trouble since its installation with the two exceptions given below.

2.6.1 Tuning Issues

It was soon discovered that the position of the cell within the NMR spectrometer probe has a significant effect on the tuning of the NMR spectrometer probe. By rotating the cell it is possible to shift the tuning frequency even to the point that it is unable to be restored to the correct position by the tuning capacitors of the NMR probe. The rotational position within the NMR spectrometer over which it is possible to tune the spectrometer is relatively small and can change after disassembly and reassembly of the cell. The position of the cell, therefore, has to be adjusted occasionally after the insertion of the sample to allow for proper tuning. It is believed that this is a result of the non-uniform metallic cell manifold interfering with the RF signal. Figure 2.11 shows the cross section of the manifold base. The two flat surfaces allow for the use of tools for the cell assembly. After each

assembly, the orientation of the manifold base can differ slightly in orientation relative to the top section of the manifold. It is this change in position which leads to the changes in the cell orientation within the NMR spectrometer after sample insertion. An attempt was made to solve this problem by exchanging the aluminium cell manifold with a titanium equivalent. There was no noticeable difference between the two manifolds.

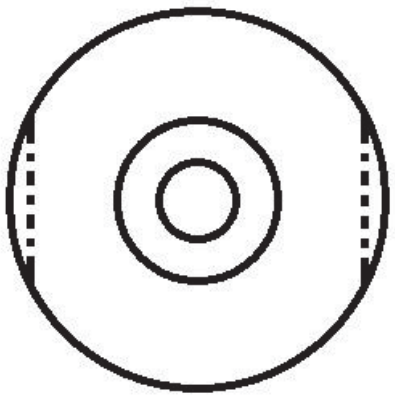


Figure 2.11: Cross Section of Cell Manifold Base as Seen From Below. The dashed lines indicate the flattened surfaces which allow for a spanner to grip this section for cell assembly.

2.6.2 Cell Failure

Unfortunately, during the course of this research, one high-pressure NMR tube suffered a critical failure. The conditions of this failure are of great importance so as to prevent this event happening again. At the time of the failure the tube was immersed in a 100 °C oil bath while at 150 MPa (60% of the rated maximum). The cell tube was immersed to a depth of approximately 7 cm for 16 hours. During this time the tube failed catastrophically, destroying the lower section of the tube (Figure 2.12). The pieces of the tube were contained in the oil bath and there was no further damage to the surrounding area. It appears that the most likely cause of this failure is the continual heating and cooling of the tube. The tube had previously been under these conditions in excess of 1000 hours, which may have caused damage to the tube's ceramic structure. The temperature difference at the oil-air interface can be eliminated as a factor. The highest point of the breakage (a fracture) was 46 mm from the base of the tube, 24 mm below the oil surface.

While the failure of the tube was undesirable it was not a result of misuse or reckless handling of the high-pressure NMR cell. These conditions were within the operational limits. All safety measures put in place to prevent further issues (harm to personnel, destruction of oil bath) worked as they were supposed to. This event can be considered an isolated incident as Daedalus Innovations reported that this was the first tube failure that they had encountered outside of testing in their lab.

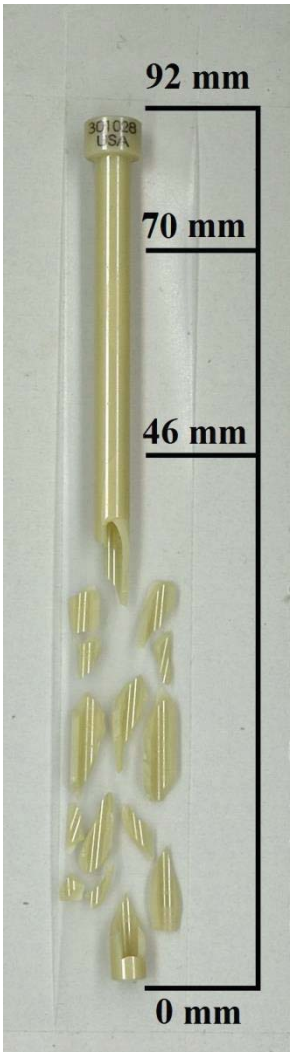


Figure 2.12: Broken High-Pressure Tube. Base of tube, 0 mm. The highest point of the breakage, 46 mm. Oil level when immersed, 70 mm. Top of tube, 92 mm.

Chapter 3 - Cytosine Deamination

3.1 Introduction

For an environment to be suitable for the development of the first life forms, it must first be capable of providing and enabling the continued existence of the primary building blocks of life. In the instances of DNA and RNA, the chemical stability of the nucleobases will be an important factor influencing the existence of self-replicating systems. In the cases where these bases are not stable, their decay will lead to errors in replication as well as the loss of catalytic functionality. The effects of temperature and solution conditions (pH, salts, concentration) on the chemical stability of these bases has been studied in the past; however, nothing is known of the effects of high pressures on their chemical stability, particularly that of cytosine which is the least stable of all the bases. Whether this is a positive or negative effect and of what magnitude, and whether high pressures have the possibility of sufficiently offsetting the negative effects of temperature, is still unknown.

3.1.1 Earlier Studies

A pioneering study by Levy and Miller [27] examined the temperature dependence of the chemical stability of individual RNA and DNA nucleobases. Their studies assessed the rate of degradation of the major nucleobases, adenine (A), cytosine (C), guanine (G), thymine (T) and uracil (U) (Figure 3.1), along with the alternative bases, xanthine, hypoxanthine and diaminopyrimidine, at temperatures ranging from 273.2 K to 373.2 K in an aqueous solution and up to 623.2 K as neat material.

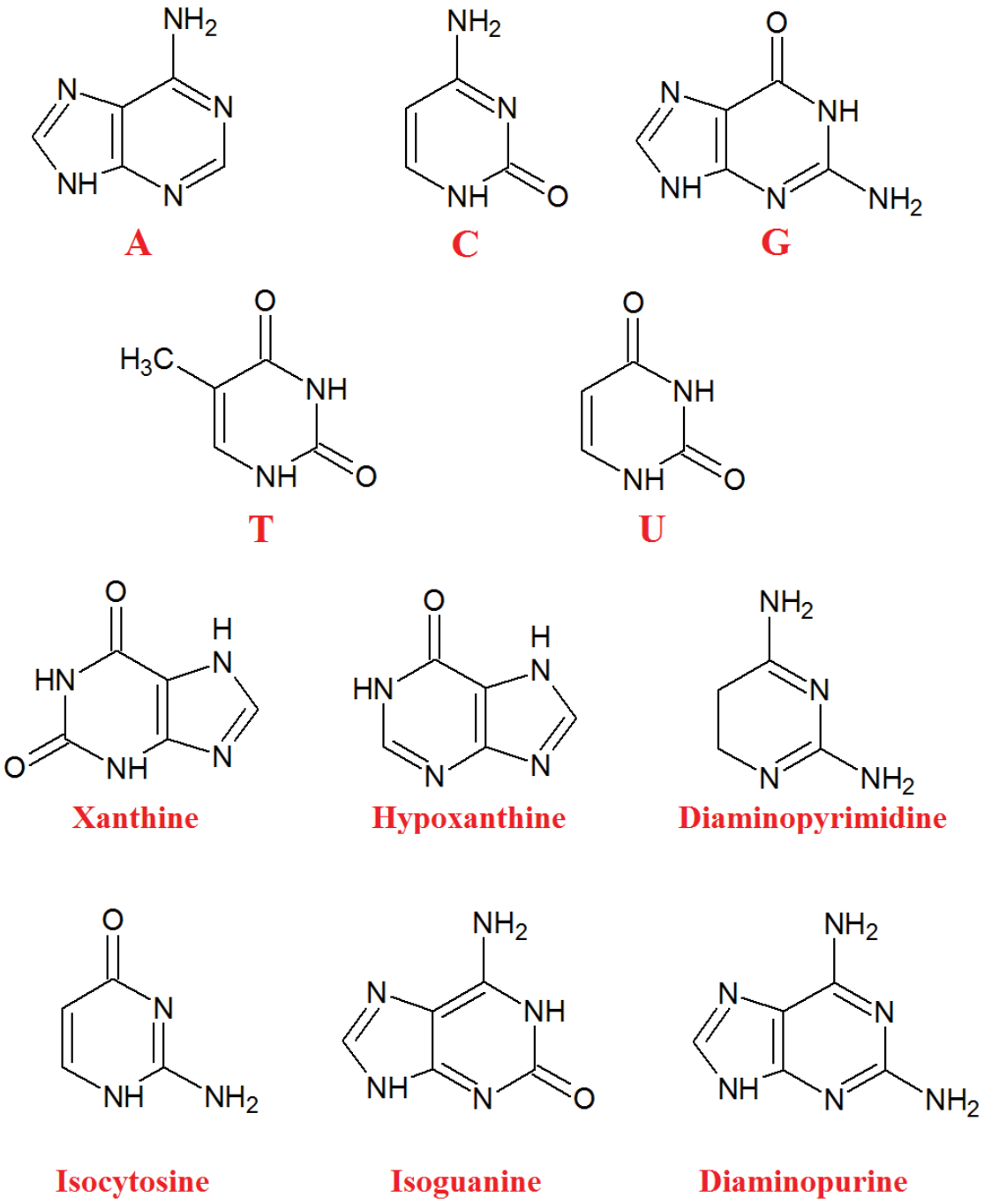


Figure 3.1: DNA/RNA Bases. Adenine (A), cytosine (C), guanine (G), thymine (T) and uracil (U). The alternate bases xanthine, hypoxanthine, diaminopyrimidine, isocytosine, isoguanine and diaminopurine are also shown.

It was discovered that, at temperatures of 623.2 K, the half-lives for decay of most bases in neat form were between 2 and 15 sec, and at 583.2 K they were between 1 and 35 min. These rates are incredibly fast, making it impossible for any significant

accumulation of these molecules to occur. With this rate of decay, a heterotrophic origin of life at ultrahigh temperatures would be highly unlikely.

At 373.2 K in aqueous solution, the longest observed decay half-life for the RNA/DNA nucleobases was for thymine with a half-life of 56 years. The half-lives for adenine, guanine and uracil are 1.0, 0.8 and 12 years respectively. At this temperature the half-life of hydrolysis cytosine is significantly shorter, only 19 days. While these half-lives are much longer than those at higher temperatures, they are still relatively short in a geological timescale. For the consideration of an origin of life at 373.2 K, particularly one involving cytosine, these compounds will have to be utilised almost immediately after their formation, otherwise this scenario would also be unlikely.

Results at 298.2 K gave the half-lives of decay for adenine and guanine to be approximately 10,000 years, whereas the half-life of hydrolysis for cytosine is only 340 years. At 273.2 K the half-lives of decay of adenine, guanine, thymine and uracil are 6×10^5, 1.3×10^6, 2×10^9 and 3.8×10^8 years, respectively. However, for cytosine, the half-life of hydrolysis at 273.2 K is only 17,000 years, which is still relatively short on the geological time scale. With the exception of cytosine, these rates of degradation at 273.2 K are approaching the rate of the destruction of organic matter as it passes through hydrothermal vents, which is considered to be the major limiting factor in organic molecule build-up [52].

3.1.2 Consequences

The results of Levy and Miller [27] showed that adenine, guanine, thymine and uracil are sufficiently stable at temperatures below 298.2 K to be available for the origin of life. However, the significantly higher rate of hydrolysis of cytosine at both high and low temperatures causes serious problems for high-temperature origin-of-life theories. Assuming that earliest life forms were rather inefficient and not adapted and stripped down for speedy turnover, in the way many extant Prokarya, Crenarchaea and Archaea are so adapted, the lack of availability of raw material for the replication of these life forms rendered unlikely a hot-start

scenario for RNA-based life forms, particularly those involving cytosine. While there have been multiple theories that attempt to compensate for this lack of stability (alternate bases [53], two-base sequences [54]), the consequences of high pressures are still unknown.

3.1.3 High-Pressure Effects

Cytosine has been recognised as the key problem in any high-temperature origin-of-life theory due to its relatively high rate of hydrolysis at temperatures above 298.2 K. What has yet to be addressed is the effect of high pressures, and whether, or not, high pressure will counteract or exacerbate the deleterious effects of high temperatures on the chemical stability of cytosine. It is believed that high pressure may have had an important role in the development of life on earth [55].

The following is a comprehensive study into the effects of high hydrostatic pressure on the chemical stability of cytosine (the individual nucleobase) and cytidine (the individual ribose nucleoside) at 373.2 K along with brief studies on the effects of pH on chemical stability under high-pressure.

3.2 Materials and Methods

3.2.1 Cytosine/Cytidine Solutions

All water used in sample preparation was milli-Q water, and masses and volumes used in preparation of all stock solutions are given in Table 3.1.

Cytosine and cytidine solutions for hydrolysis measurements were prepared as follows:

Stock Solution	Mass used /g	Volume used /mL	Concentration /M	Molar Mass /gmol^{-1}	Supplier
Cytosine	0.0222	10.00	0.0200	111.1	Sigma
Cytidine	0.0486	10.00	0.0200	243.2	Sigma
TMSP-d4	0.0073	10.00	0.0050	156.26	Merck
Sodium azide	0.0013	10.00	0.0020	65.01	Merck
NaH_2PO_4. $2H_2O$	1.5601	50.0	0.200	156.01	Riedel-de Haën
Na_2HPO_4. $2H_2O$	1.7799	50.0	0.200	177.99	Riedel-de Haën
NaCl (pH 6)	0.2600	5.00	0.889	58.5	Merck
NaCl (pH 7)	0.1626	5.00	0.556	58.5	Merck
NaCl (pH 8)	0.1298	5.00	0.365	58.5	Merck

Table 3.1: Table of Stock Solutions for Cytosine/Cytidine Experiments. Each entry shows concentration, mass of compound used, volume made (in milli-Q water) and supplier. The pH value given with the NaCl solutions indicate the pH at which the solution is to be used.

Solutions for both cytosine and cytidine were prepared by combining 2.5 mL of 20 mM stock solutions of either cytosine or cytidine, separately, with 1 mL of 5.0 mM 3-(trimethylsilyl)-2,2',3,3'-tetradeuteropropionic acid (TMSP-d4) stock solution and 0.5 mL of sodium azide stock. The solution was then made up to 5 mL with water.

Several 0.20 M phosphate buffer solutions were also prepared by adding 0.20 M disodium hydrogen phosphate stock solution to 0.20 M monosodium dihydrogen phosphate stock solution. The disodium hydrogen phosphate was added in the presence of a pH meter to ensure the correct pH. The pH values of each buffer solution are such that they will be at the desired pH when under high-pressure conditions as given in Table 3.2.

Solutions for the hydrolysis measurements for cytosine/cytidine solutions were made by combining 0.1 mL of the cytosine/cytidine solution prepared previously with 0.5 mL of the desired buffer, 0.15 mL of NaCl solution (0.889 M for pH 6

Pressure	pH under pressure	pH shift	Corrected pH at 0.1 MPa
0.1	6.00	0.00	6.00
0.1	7.00	0.00	7.00
0.1	8.00	0.00	8.00
50	7.00	0.20	7.20
100	7.00	0.39	7.39
150	6.00	0.56	6.56
150	7.00	0.56	7.56
150	8.00	0.56	8.56
200	7.00	0.71	7.71

Table 3.2: Buffer pH Corrections. This table displays the buffer solutions used for these studies. The first two columns show the intended pressure conditions and the pH required. The third column shows the resulting pH shift and the fourth column shows the pH value at 0.1 MPa required to achieve the correct pH under pressure.

0.556 M for pH 7 and 0.365 for pH 8), 0.25 mL water and 0.1 mL D_2O (MagniSolv). This gave a final solution of 1.00 mM cytosine, 50 mM phosphate buffer, 0.02 mM sodium azide and 0.2 mM TMSP-d4 in H_2O/D_2O (90:10 v/v).

The addition of the NaCl solution adjusted the ionic strength of all solutions to 0.20 M. Note that this is not the case when the solutions are at atmospheric pressure when the pH levels are at their initial values. When the solutions are pressurised at the desired level, the shift in the pH alters the ions present in solution and brings the ionic strength back to 0.20 M. Concentrations of the NaCl solutions were determined from the following equations.

The ionic strength, I, of a solution is given by:

$$I = \frac{1}{2}\sum_{i=1}^{n} c_i Z_i^2 \tag{3}$$

Where c_i is the molarity of ion i and Z_i is its charge.

Inserting the relevant ions into this equation results in:

$$I = \frac{1}{2}\left([\mathrm{Na^+}]_s + [\mathrm{Cl^-}]_s + 4[\mathrm{HPO_4^{2-}}] + [\mathrm{Na^+}]_{di} + [\mathrm{H_2PO_4^-}]\right) + [\mathrm{Na^+}]_{mono} \quad (4)$$

Since $[NaCl] = [Na^+]_s = [Cl^-]_s$, $[HPO_4^{2-}] = 2[Na^+]_{di}$ and $[HPO_4^{2-}] = [Na^+]_{mono}$, equation (4) can be simplified to:

$$I = [\mathrm{NaCl}] + 3[\mathrm{HPO_4^{2-}}] + [\mathrm{H_2PO_4^-}] \quad (5)$$

To define the concentration of the buffer ions the Henderson-Hasselbalch equation must first be considered. This defines pH as:

$$\mathrm{pH} = \mathrm{p}K_a + \log\left(\frac{[\mathrm{HPO_4^{2-}}]}{[\mathrm{H_2PO_4^-}]}\right) \quad (6)$$

This can be rearranged to give the ratio of conjugate acid and base, *x*, as:

$$\left(\frac{[\mathrm{HPO_4^{2-}}]}{[\mathrm{H_2PO_4^-}]}\right) = 10^{\mathrm{pH}-\mathrm{p}K_a} = x \quad (7)$$

By defining the concentration of the buffer, *C*, as the sum of the concentrations of the two buffer ions, equations (7) and (8) can be combined to obtain expressions for concentration of the ions in terms of *C* and *x*.

$$C = [\mathrm{HPO_4^{2-}}] + [\mathrm{H_2PO_4^-}] \quad (8)$$

$$[\mathrm{H_2PO_4^-}] = \frac{C}{1+x} \quad (9)$$

$$[\mathrm{HPO_4^{2-}}] = \frac{Cx}{1+x} \quad (10)$$

Equations (9) and (10) can be inserted into equation (4) to obtain an expression for the ionic strength in terms of C and x

$$I = [\mathrm{NaCl}] + \frac{C}{1+x}(3x+1) \quad (11)$$

Lastly, equation (11) can be arranged to give an expression for the NaCl concentration required.

$$[\mathrm{NaCl}] = I - \frac{C}{1+x}(3x+1) \tag{12}$$

NaCl concentrations used are shown below in Table 3.3.

Buffer pH	Final salt concentration /mol L^{-1}	Initial salt concentration before dilution /mol L^{-1}
6.00	0.133	0.889
7.00	0.083	0.556
8.00	0.055	0.365

Table 3.3: Salt Concentrations. This table shows the salt concentrations required to achieve an ionic strength of 0.2 M for the solutions of different pH. The concentration of the initial solutions before dilution with the cytosine/cytidine solution is also shown.

3.2.2 Incubation

The samples were incubated individually as follows.

A sample was first inserted into the high-pressure NMR cell, which was then sealed and set to the desired pressure. An initial ^{1}H NMR spectrum was then recorded. The high-pressure NMR cell containing the sample was then carefully placed in an oil bath at a depth of approximately 7 cm (Figure 3.2) and heated to 100 °C. The sample was incubated for a period of no less than 135 hours.

During incubation, the cell was removed periodically to allow for ^{1}H NMR spectra to be recorded. This involved removing the cell from the oil bath where it was cooled to room temperature and cleaned to remove any oil. The cell was then inserted into the NMR spectrometer to record a ^{1}H NMR spectrum. During this time (approximately 45 minutes), the oil bath was allowed to cool to a temperature of approximately 333 K. This ensured that the high-pressure cell was not inserted directly into the oil at 373.2 K. After the recording of measurements, the cell was then removed from the spectrometer and returned to the oil bath and reheated to

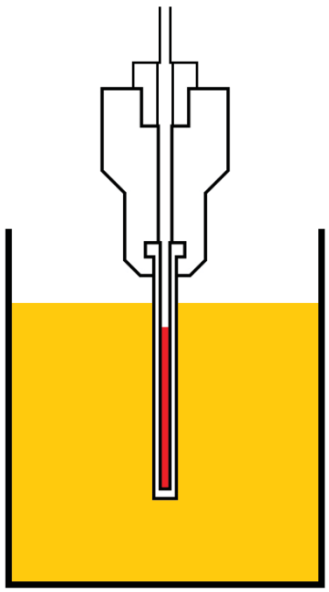


Figure 3.2: High-Pressure NMR Cell in Oil Bath. This figure shows the depth to which the high-pressure NMR cell was placed in the oil bath. It can be seen that the entire sample (red) within the cell is below the oil level.

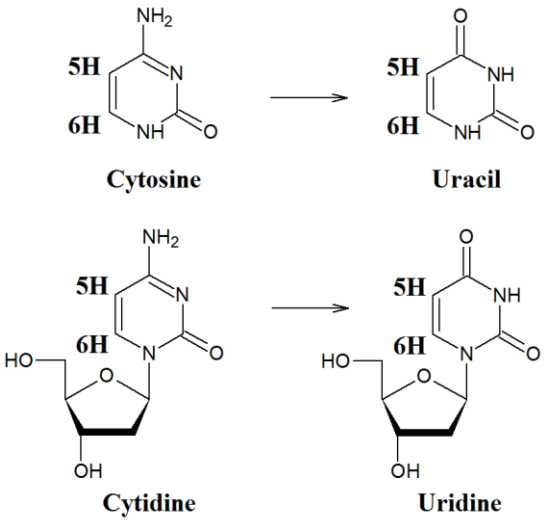
373.2 K. No significant change in pressure resulting from the temperature changes was observed during cooling or heating.

3.2.3 Hydrolysis Measurements

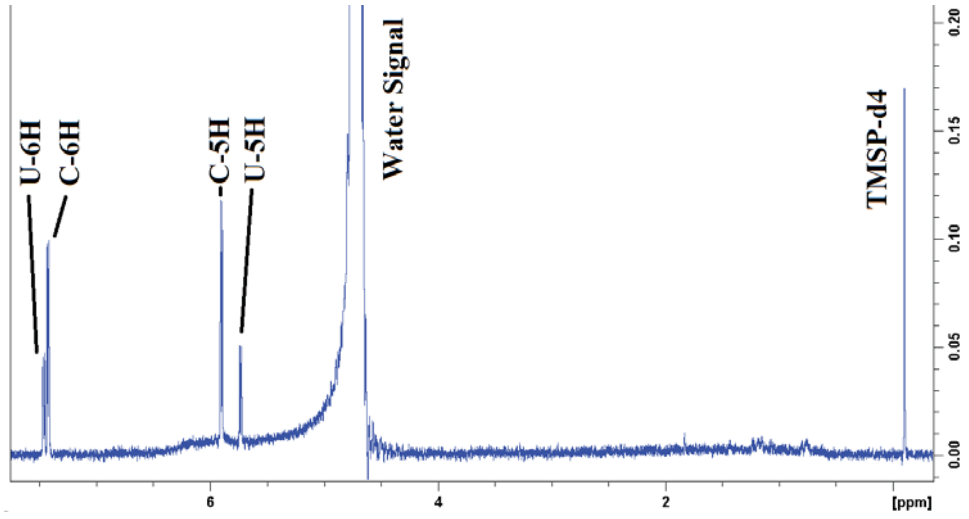
Hydrolysis of cytosine and cytidine was monitored by ^{1}H NMR spectroscopy. Spectra were obtained using a Bruker Avance 500 MHz NMR spectrometer. The spectrometer was fitted with a TXI 5 mm ^{1}H-detection inverse probe with a *z*-field gradient coil. All spectra were recorded with 64 scans with a relaxation time of 10 s, spectral width 20.7 ppm and 65536 points. A pre-saturation pulse was used to suppress the dominant water peak in the spectrum. A shaped pulse was initially used for improved water suppression but was soon abandoned in favour of the pre-saturation pulse. This was due to the shaped pulse interfering with the values of the integrals of several proton peaks in the NMR spectrum.

For the hydrolysis of cytosine to uracil, the integrals of the proton peaks corresponding to protons 5H and 6H (Figure 2.1) for both cytosine and uracil were

measured for each NMR spectrum using the integral of the TMSP-d4 peak (the NMR calibration peak of a known concentration which remained constant throughout the hydrolysis reaction) as a reference. The ratio of the integrals of both the cytosine and the uracil to the integral of the TMSP-d4 reference peak were used to determine what proportion of the initial cytosine had been hydrolysed to uracil and what proportion of cytosine remained. This same method was used to determine the level of cytidine hydrolysis to uridine.



a)



b)

Figure 3.3: Cytosine/Cytidine Hydrolysis. a) Cytosine to uracil and b) cytidine to uridine indicating hydrogens 5H and 6H. c) An example spectrum of cytosine during hydrolysis. The peaks from hydrogens 5H and 6H from both cytosine and uracil have been labelled: for cytosine, C5 and C6, and for uracil, U5 and U6. These peaks were used to determine the rate of hydrolysis.

3.3 Results

^{1}H NMR spectra were recorded for the hydrolysis of both cytosine and cytidine at 373.2 K and at pressures ranging from 0.10 MPa to 200 MPa for multiple time points over at least 135 hours. The ^{1}H NMR spectra were used to determine the level of hydrolysis for each of these time points.

The NMR spectra, especially of cytidine, were closely inspected for evidence of decomposition additional to the deamination to uracil. No evidence of decomposition was found by way of peaks not assignable to cytosine/uracil or cytidine/uridine. Moreover, the values of the integrals of the 6H and 5H proton signals for uracil and cytosine (or cytidine), normalised to the integral of the TMSP -d4 reference signal, were constant within 4.0% over the time course of the data gathering.

The spontaneous hydrolysis of cytosine or cytidine is given in Figure 3.4 [56, 57].

The rate of this reaction is determined by the slow substitution of the NH_2 group. Considering this, the rate equation can be written as

$$Rate = k\,[\text{cytosine}][H_2O] \tag{13}$$

here k is the rate constant of the reaction. Since the concentration of H_2O is constant in an aqueous solution, the rate equation can be rewritten as

$$Rate = k_{obs}\,[\text{cytosine}] \tag{14}$$

where k_{obs} is the observed rate constant. The same will also be true for cytidine. Since this is a first-order rate equation, the first-order integrated rate equation can be used to give the relationship between the concentration of the nucleobase (cytosine/cytidine) and time, *t*.

$$[n\text{ucleobase}] = [\text{nucleobase}]_0 e^{-k_{obs}t} \tag{15}$$

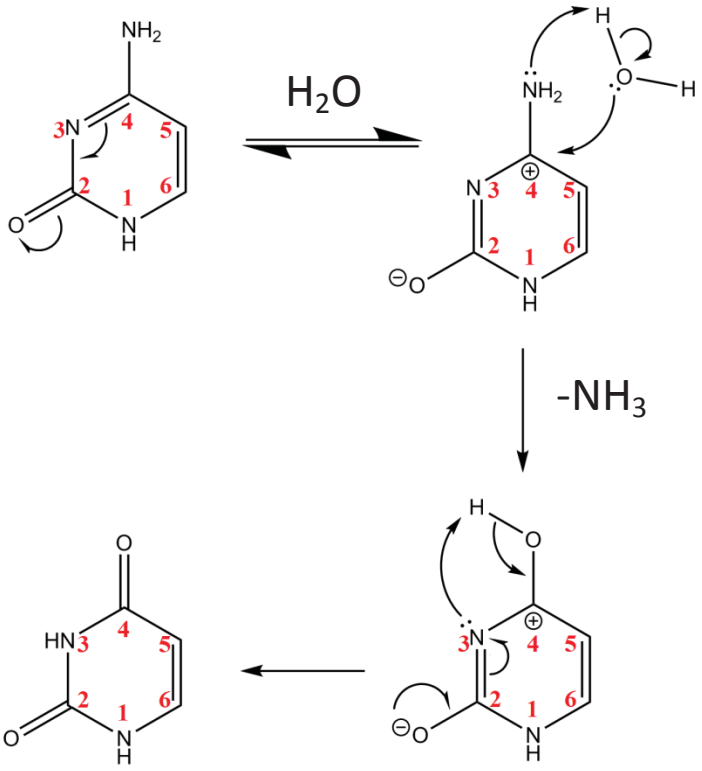


Figure 3.4: Mechanism of Cytosine Deamination. This figure outlines the mechanism by which cytosine or cytidine undergoes deamination. A resonance shift leaves a positive charge on C4 making it susceptible to attack from a water molecule. The amino group then leaves as an NH_3.The N3 then takes the proton from the neighbouring OH. A final resonance shift restores the molecule to its final state as uracil. Alternative mechanisms have been proposed including catalytic activity from the buffer [58]. However, these mechanisms have not been confirmed and still share the same rate-limiting step. As such, the mechanism outlined in this figure will suffice for understanding the rate relationship.

where [nucleobase] is the concentration of the nucleobase at time t, $[\text{nucleobase}]_0$ is the initial concentration of the nucleobase and k is the rate constant. This can be rearranged to give the fraction of the remaining nucleobase (N_f) as a function of time

$$N_f = \frac{[\text{nucleobase}]}{[\text{nucleobase}]_0} = e^{-k_{obs}t} \tag{16}$$

Finally, this can be rearranged to give a linear relationship

$$\ln N_f = -k_{obs}\, t \tag{17}$$

A plot of ln N_f versus time can be used to determine the rate constant for the hydrolysis of the nucleobase.

The relationship between the observed rate constant, k_{obs}, and pressure, P, is given by equation (18) [49]

$$\frac{\partial \ln k_{obs}}{\partial P} = -\frac{\Delta V^{\ddagger}}{RT} \qquad (18)$$

where R is the gas or universal constant and T is the absolute temperature in Kelvin. $\Delta V^{\ddagger}$ is the reaction volume of the reaction and can be determined from a plot of ln k_{obs} versus pressure.

3.3.1 Cytosine Hydrolysis

Hydrolysis reactions of cytosine were performed under multiple pressure and pH incubation conditions. In each instance, a minimum of seven hydrolysis measurements was made over an approximately one-week incubation period. Figure 3.5 shows ln N_f versus time for each pressure at pH 7. These conditions along with the rate constants and associated estimated standard deviations derived from the gradients of the resulting lines are tabulated in Table 3.4.

A brief examination of the pH dependence of cytosine hydrolysis was carried out at pressures of 0.1 and 150 MPa (Figure 3.6). It can be seen in both curves that the rate of hydrolysis is the lowest in the region of pH 7. This result is in general agreement with the results seen by Garrett [59]. This result also shows that there is a reasonably strong dependence of pH on the rate constant k_{obs}, which may affect the outcome of the hydrolysis measurements. This reinforces the importance of correcting for pressure effects on the pH of the buffer.

It is clear from Table 3.4 that the observed rate constant increases monotonically as the pressure increases. At 0.1 MPa the rate constant is (0.42 ± 0.04) x 10^{-6} s^{-1} increasing to (0.88 ± 0.04) x 10^{-6} s^{-1} at 200 MPa. The pressure profile of the rate constants for cytosine at 373.2 K is shown in Figure 3.7, as a plot of ln k_{obs} versus pressure. The resulting gradient has a value of 3.8 x 10^{-9} Pa^{-1}. Using equation (18), $\Delta V^{\ddagger}$ has been calculated to be -11.7 ± 1.2 cm^3 mol^{-1}. This result is similar in magnitude to the molar volume of water, 18.01 cm^3 mol^{-1}.

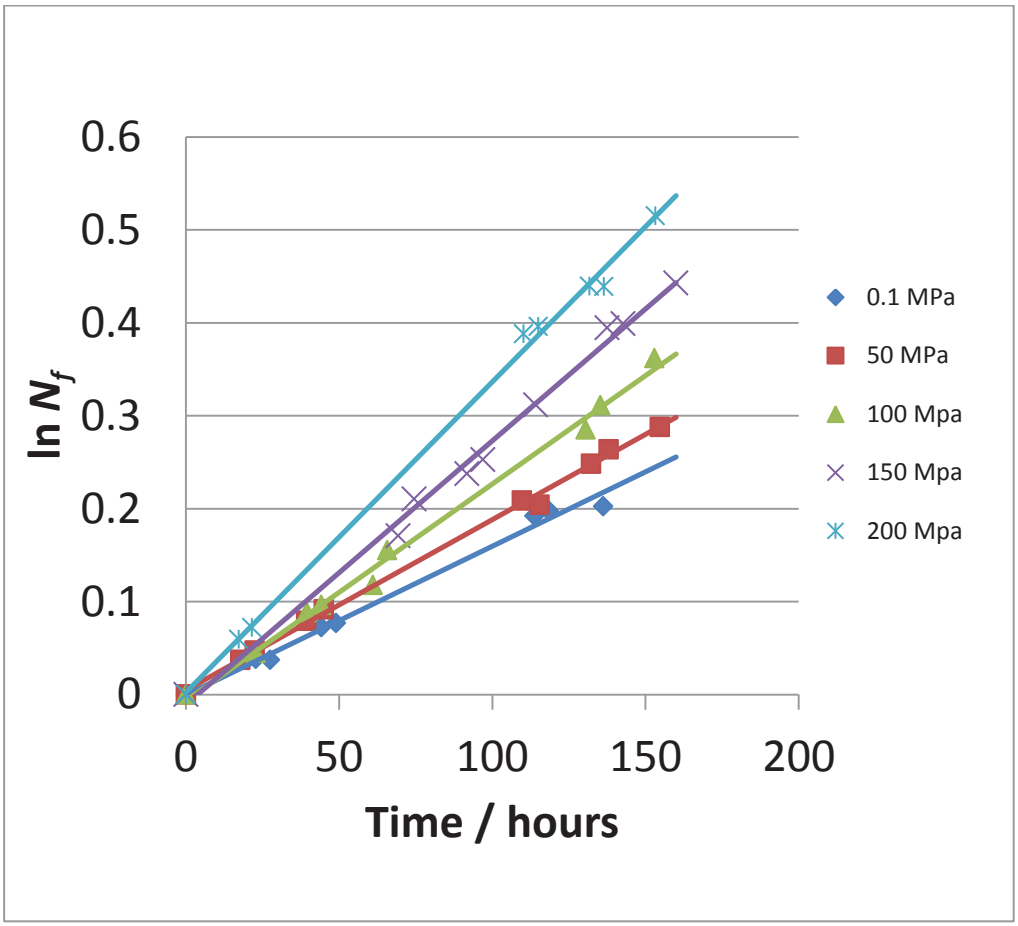


Figure 3.5: ln N_f versus Time for Cytosine Hydrolysis at pH 7. The results of each pressure condition are shown.

	Pressure / MPa				
	0.1	**50**	**100**	**150**	**200**
	k / x10^{-6} s^{-1} [$t_{1/2}$ / day]	k / x10^{-6} s^{-1} [$t_{1/2}$ / day]	k / x10^{-6} s^{-1} [$t_{1/2}$ / day]	k / x10^{-6} s^{-1} [$t_{1/2}$ / day]	k / x10^{-6} s^{-1} [$t_{1/2}$ / day]
Cytosine pH 6.0	0.59 ± 0.07 [13.6]			0.87 ± 0.07 [9.3]	
Cytosine pH 7.0	0.42 ± 0.04 [19.3]	0.49 ± 0.04 [16.5]	0.62 ± 0.03 [13.1]	0.72 ± 0.06 [11.1]	0.88 ± 0.04 [9.1]
Cytosine pH 8.0	0.43 ± 0.03 [18.9]			0.78 ± 0.04 [10.3]	

Table 3.4: Conditions for Incubation of Cytosine. This table displays the calculated k_{obs} values for the hydrolysis of cytosine for each pressure and pH condition at 373.2 K. The half-life ($t_{1/2}$) in days is also shown below each k_{obs} value.

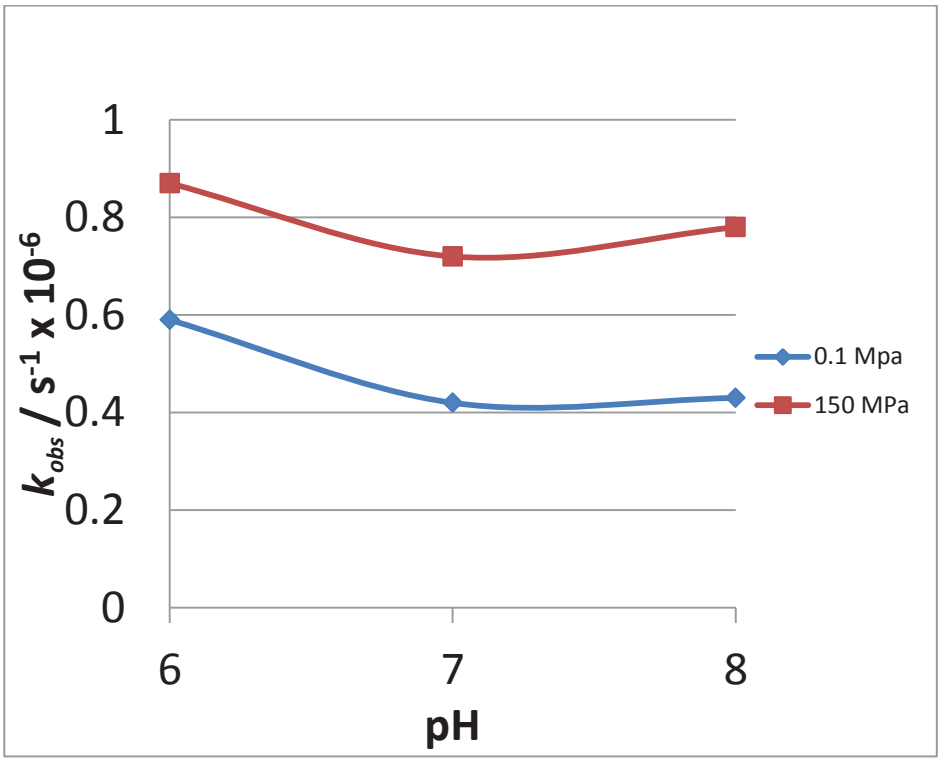


Figure 3.6: k_{obs} versus pH. k_{obs} values at pH 6.0, 7.0 and 8.0 for both 0.1 MPa and 150 MPa. Connecting lines are as a visual reference only.

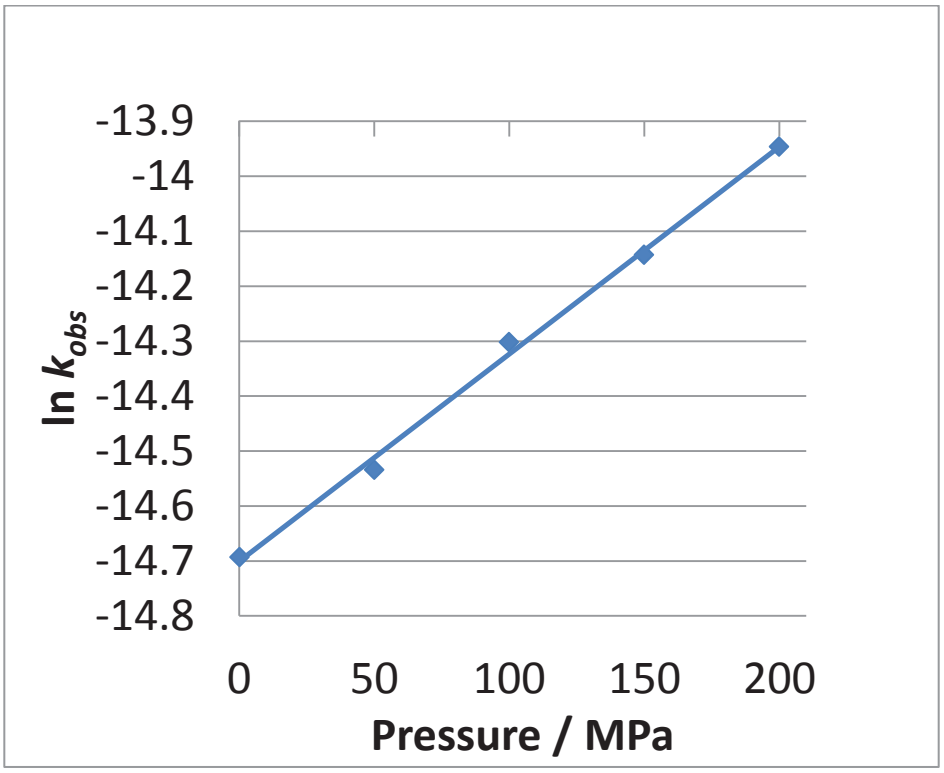


Figure 3.7: ln k_{obs} versus Pressure for Cytosine. From the gradient of this plot, the reaction volume has been calculated to be -11.7 ± 1.2 cm^3 mol^{-1}.

3.3.2 Cytidine Hydrolysis

Hydrolysis reactions of cytidine were performed at pressures of 0.1 MPa and 150 MPa. In each instance, a minimum of seven hydrolysis measurements were made over an approximately one-week period. Figure 3.8 shows ln N_f versus time for both pressures. The rate constants and associated estimated standard deviations derived from the gradients of the resulting lines are tabulated in Table 3.5.

The hydrolysis of cytidine shows similar behaviour to the hydrolysis of cytosine. However, the rate of hydrolysis of cytidine is clearly faster than that of cytosine at both low and high pressures. The observed rate constants for cytidine at 0.10 MPa and 150 MPa at 373.2 K are (0.44 ± 0.04) x 10^{-6} s^{-1} and (0.87 ± 0.08) x 10^{-6} s^{-1} respectively.

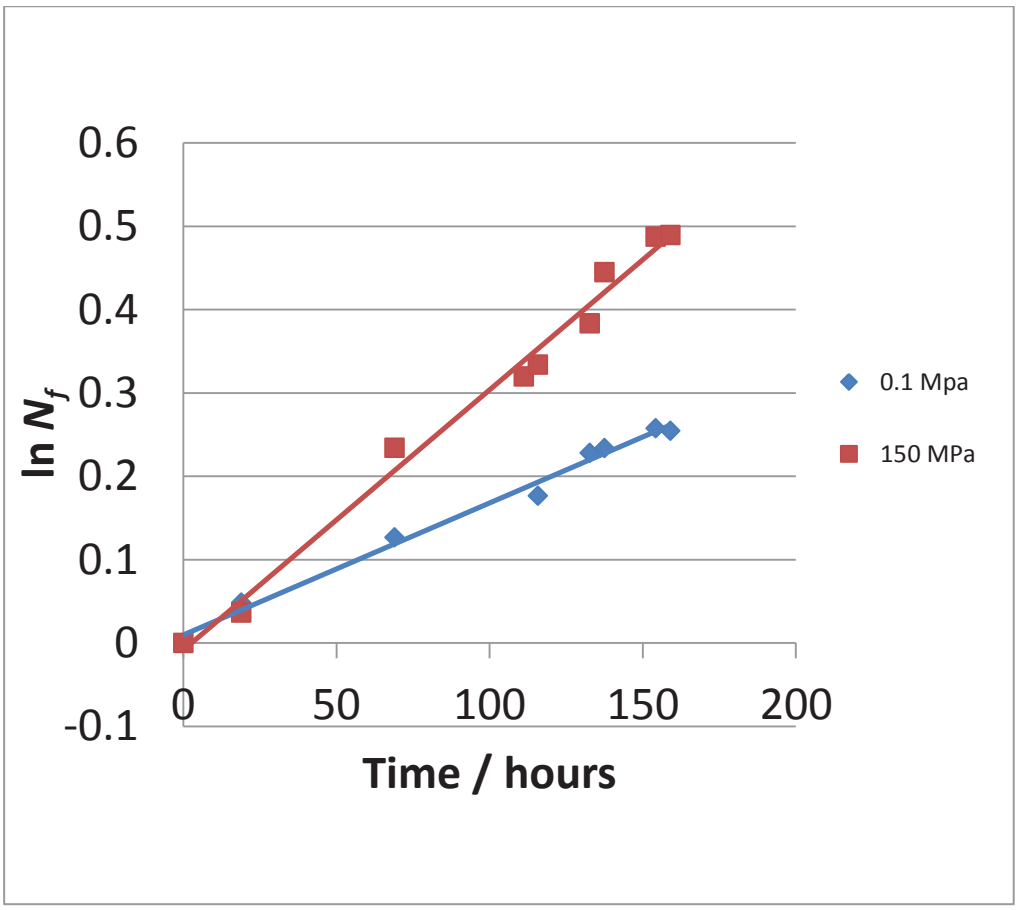


Figure 3.8: ln N_f versus Time for Cytidine Hydrolysis at pH 7.

	Pressure / MPa	
	0.1	**150**
	k / x10^{-6} s^{-1} [$t_{1/2}$ / day]	k / x10^{-6} s^{-1} [$t_{1/2}$ / day]
Cytidine pH 7.0	0.44 ± 0.04 [18.2]	0.87 ± 0.08 [9.3]

Table 3.5: Conditions for Incubation of Cytidine. This table displays the calculated k_{obs} values for the hydrolysis of cytidine at 0.1 MPa and 150 MPa at pH 7. The half-life ($t_{1/2}$) in days is also shown below each k_{obs} value.

The gradient from a plot of ln k_{obs} versus pressure was found to be 4.5 x 10^{-9} Pa^{-1}. With the use of equation (18), $\Delta V^{\ddagger}$ has been calculated to be -14.6 ± 1.9 $cm^3\,mol^{-1}$ (error determined from the least-squares of the error in the two data points). This value of the reaction volume for cytidine is significantly larger than the result for cytosine of -11.7 ± 1.2 $cm^3\,mol^{-1}$.

3.4 Discussion

The relatively quick rate of hydrolysis of cytosine to uracil at 373.2 K and ambient pressure, compared to geological time scales, has previously been cited as a limiting factor for the evolution of life at high temperatures [27]. With regard to all the possible locations for the development of life, there is a wide range of different conditions to consider, from acidic, high-pressure, high-temperature deep-sea black smokers to alkaline, low-pressure, and moderate-temperature haline springs. Because of this, studies were undertaken to examine the effect of pressure and pH on cytosine and cytidine hydrolysis to determine what effect pressure has on the chemical stability of cytosine and whether there is any possibility of high pressures offsetting the negative effects of high temperatures on nucleobases chemical stability in neutral pH conditions.

Somewhat counter to expectations, the rate of hydrolysis of cytosine **increases** with pressure, such that at 200 MPa the rate constant for deamination at 373.2 K is doubled compared to that measured at an ambient pressure of 0.1 MPa. The

observed rate constant for the hydrolysis of cytosine at 0.1 MPa and 373.2 K was found to be (0.42 ± 0.04) x 10^{-6} s^{-1}. This compares well with that determined by Levy and Miller [27] of 0.427 x 10^{-6} s^{-1} for cytosine under identical conditions.

It has been observed that under the same conditions, the rate of hydrolysis of cytidine is faster than that of cytosine, the difference being more pronounced and significant at higher pressures. Respectively, the observed rate constants for the hydrolysis at 373.2 K for cytosine and cytidine are at 0.1 MPa: (0.42 ± 0.04) x 10^{-6} s^{-1} and (0.44 ± 0.04) x 10^{-6} s^{-1}, and at 150 MPa, (0.72 ± 0.06) x 10^{-6} s^{-1} and (0.87 ± 0.08) x 10^{-6} s^{-1}. This shows increases in the rate of hydrolysis compared to the atmospheric value of 5% to 20%, the increase at 150 MPa being marginally significant.

From the plot of ln k_{obs} versus pressure, the activation volume, $\Delta V^{\ddagger}$, for the hydrolysis of cytosine and cytidine at 373.2 K was calculated to be -11.7 ± 1.2 cm^3 mol^{-1} and -14.6 ± 1.4 cm^3 mol^{-1}. Respectively, these results appear to be typical for a biochemical process, where the magnitude of values for $\Delta V^{\ddagger}$ are within the range of 0 to 50 cm^3 mol^{-1} [60] . The negative value is a consequence of an increased rate of deamination with increasing pressure. The difference between the two results for cytosine and cytidine indicate that the hydrolysis rate for cytidine, which has the attached ribose ring, is more susceptible to the effects of pressure.

It is important to note that there was a small degree of estimation during the process of measuring the integrals of a small number of the proton peaks. This estimation is a result of the partial overlap of the peaks of interest (Figure 3.9). In this instance, while it can be easy to estimate where to divide the integral visually, the chosen mid-point between the peaks may not necessarily be accurate.

To draw comparisons to high-pressure conditions relevant to the deep-sea environments on earth, the derived reaction volume has been used, along with equation (18), to back calculate the rate constants for hydrolysis of cytosine at three significant ocean depths. These can be compared to the observed rate constant for the hydrolysis of cytosine at 373.2 K is (0.42 ± 0.04) x 10^{-6} s^{-1} at sea level (atmospheric pressure of 0.1 MPa).

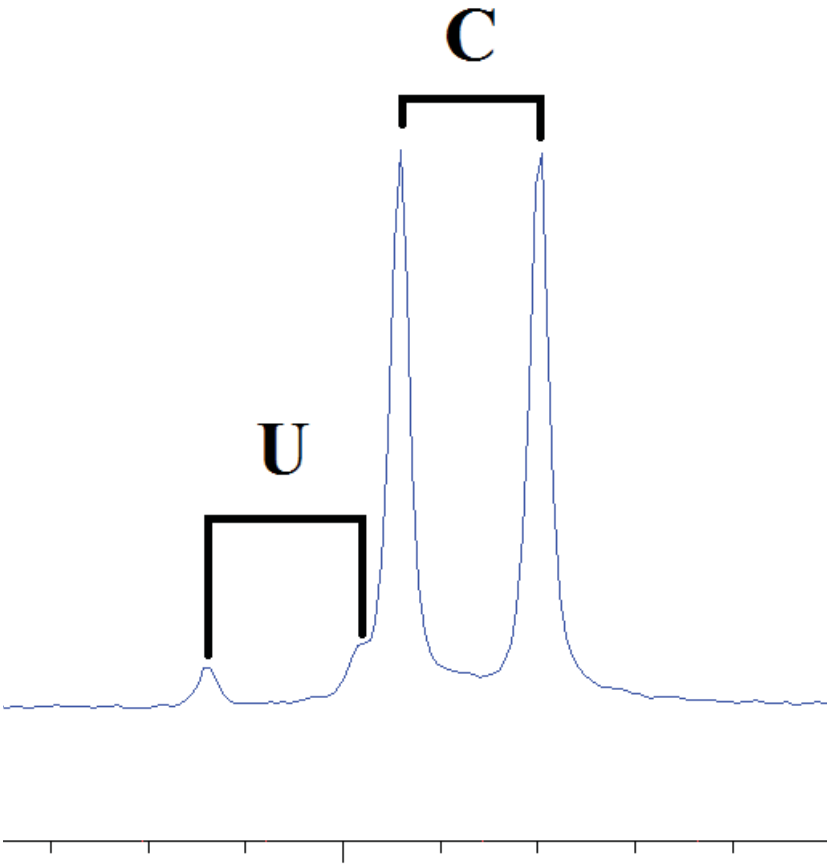


Figure 3.9: Overlapping Peaks. An example of overlapping proton peaks where the midpoint of the proton signal was estimated. In this instance the uracil signal could possibly be offset by the overlap.

The average ocean depth is 3790 m [61] and corresponds to a pressure of 38.2 MPa. At this pressure, the rate constant for the hydrolysis of cytosine has been calculated to be 0.48×10^{-6} s^{-1}. This is approximately 14% higher than the value at atmospheric pressure. While this result is larger, it is unlikely to be significant in terms of any geological time scales. The deepest submarine hydrothermal vent which has been discovered is at a depth of 4100 m [23]. At this depth the pressure is 41.3 MPa and the observed rate constant has been calculated to be 0.49×10^{-6} s^{-1}, approximately 16% higher than the value at atmospheric pressure.

The greatest ocean depth has measured at 10,920 m [62] and corresponds to a pressure 110 MPa. At these pressure conditions the rate constant for the hydrolysis of cytosine has been calculated to be 0.62×10^{-6} s^{-1}. This is 47% higher than the value at atmospheric pressure. In this instance the result is significantly higher. This increase in the rate of hydrolysis of cytosine will have a large effect on

its availability and on the stability of any biomolecules which contain cytosine. However, there is only a very small proportion (<1%) of the ocean below 10,000 m.

The examination of the pH dependence of the hydrolysis of cytosine has shown that the pH dependence of hydrolysis at high pressures is similar to that at ambient pressure. At both 0.1 MPa and 150 MPa, the plots of the observed rate constant versus pressure indicate that the pH that will result in the lowest rate of hydrolysis is located somewhere just above pH 7. At atmospheric pressure this result is to be expected since the nucleobases are most stable at pH 7 [59]. The similar result under high-pressure conditions shows that this trend is still observable at high pressures. Therefore, pH 7 is an ideal pH for minimal cytosine hydrolysis.

The instability of biomolecules has long been recognised as a weakness in the argument for a hot-start origin-of-life theory [63]. Cytosine has been observed to be the least stable of all the nucleobases having a half-life of 340 yr at 298.2 K (compared to ~ 10,000 yr for adenine and guanine at 298.2 K) but having a half-life of only 19 days at 373.2 K (compared to ~ 1 yr for adenine and guanine). Even at 298.2 K, the rate of deamination of cytosine is very fast on the geological time scale and therefore decreases the likelihood of a high-temperature origin of life involving cytosine. Put another way, for a single-stranded, 4000-base polynucleotide containing 1000 cytosine bases, in the absence of other stabilising factors, at 298.2 K, the rate of degradation would result in an initial decay rate of approximately one cytosine every 0.5 years. With no way to repair these changes, this will result in a rapid rate of decay of the original sequence.

It had been argued that the instability of RNA/DNA and its components at high temperatures may be offset by high pressures in a high-temperature/high-pressure theory [55]. Results show that this is not the case and that it is in fact the opposite for the hydrolysis of both cytosine and cytidine. This is solid evidence against a high-temperature/high-pressure origin of life involving cytosine or derivatives. There is also a possibility for an origin of life that does not involve cytosine as a nucleobase, in genetic material that is either two-base coded with just A and U or where the CG pair has been substituted for an alternate base pair

(isoguanine and isocytosine, diaminopurine and U, diaminopyrimidine and xanthine [27, 64] (Figure 3.1)). However, numerical simulations indicate that a model containing only A and U does not lead to the unique stable folded RNA structures necessary for catalytic functions [7, 65]. This would be an important limitation if proteins had yet to be established as the major group for catalytic function.

3.5 Conclusions

As previously observed, the rate of hydrolysis of cytosine at 100°C is relatively short on the geological time scale. The data obtained on the chemical stability of cytosine and cytidine under high-pressure conditions at pH 7.0 and 373.2 K indicate that both cytosine and cytidine have significantly faster rates of hydrolysis at high pressures. The data also show that the pH dependence on the rate of hydrolysis for cytosine at both 0.10 MPa and 150 MPa is approximately the same.

The decrease in the chemical stability of cytosine disfavours high-temperature/high-pressure origin-of-life theories, particularly those involving cytosine. While the rate of hydrolysis of cytosine favours a low temperature origin-of-life, it is still unknown what effects pressure will have at these low temperatures. If these low-temperature effects are similar to those observed at high temperatures, the effect of pressure on the rate of hydrolysis will still be rather weak.

3.6 Future Work

There are a large number of other conditions under which the hydrolysis of cytosine can be studied. The studies outlined in this chapter represent only a small selection of the possibilities that are available. The scope of this research was intentionally limited, primarily because of time restraints and secondarily because of stresses on equipment. However, there is a great intent to continue these studies

to increase the understanding of how different environments affect the deamination of cytosine. Other variable conditions which have been considered for possible future work are discussed below.

3.6.1 Salts

There is a possibility that different salts may affect the rate of hydrolysis differently. Studies into the effects of magnesium salts on the rate of hydrolysis have been considered. The early stages of these studies had been started; however, the failure of the high-pressure NMR cell tube lead to the discontinuation of these experiments. It was decided that further experiments involving the heating of the high-pressure cell in an oil bath should cease until a later date. There is also a desire to study how ionic strength may change how pressure affects the rate of hydrolysis.

3.6.2 pH Dependence

The effect of pH on the reaction volume of cytosine has not been fully covered in the course of this research. The current studies have only been concerned with locating an approximate pH at which there is a hydrolysis rate minimum and whether this is observed under high-pressure conditions. Studies into the effects of pH on the reaction volume of cytosine hydrolysis would be very important due to the myriad of different pH environments in nature. This would complement the work already done. The importance of this work has not gone unnoticed with time constraints and the failure of the high-pressure cell delaying the undertaking of this research.

3.6.3 Molecular Crowding and Effects of Solutes

The studies performed so far have only involved solutions with a minimal presence of other solutes (buffer reagents, NMR reference, and sodium azide). However, this does not accurately reflect the conditions found in the real world. There have been considerations as to the effects of molecular crowding and other solutes. It is believed that molecular crowding may alter the rate of hydrolysis, while other solutes, such as amino acids, may have both stabilising and destabilising effects. This would be an in-depth study with multiple conditions to assess (concentrations of solutes, different molecular crowding agents and amino acids) leading to a lengthy but worthwhile study.

Chapter 4 - Physical Stability of Hexamers

4.1 Introduction

With the possibility of high pressure being an important factor in the development of the first forms of life, the need to study the effects of high pressures on the physical stability of DNA, RNA and proteins is as important as any other environmental condition. The effects of temperature and solution conditions (pH, salts, concentration) on the physical stability of biomolecules have been studied extensively on a multitude of different molecules. This has resulted in a comprehensive wealth of knowledge on how these environmental conditions affect physical stability. By contrast, studies on the effects of high pressures on the physical stability of DNA, RNA and proteins are relatively new with very little known, particularly for DNA and RNA.

4.1.1 Effects of High Pressure on Protein Stability

With the improvement of high-pressure experimental techniques, the number of studies into the effects of high pressure on the structure and physical stability of proteins has increased dramatically. With their complex tertiary structures, the effects of high pressures on the physical stability of proteins are strongly dependent on the nature of that structure. It has been observed that high pressures (up to 500 MPa) are capable of denaturing protein structures at room temperatures [38, 44, 66, 67]. These denaturing effects are most commonly a result of the hollow structure of many proteins. Under high-pressure conditions these hollow spaces allow for a significant volumetric decrease resulting in the loss of the tertiary structure. There has also been a large number of studies performed assessing how high pressures alter the 3D structure of proteins [60, 68, 69]. The results of these studies vary from protein to protein and even within the different

sections of the proteins themselves. Williamson *et al.* [70] has observed that different sections of bovine pancreatic trypsin inhibitor (BPTI) expand or contract at 200 MPa. However, the structure of BPTI showed an overall decrease in volume, elongating slightly, with most of the structural changes occurring in the loops of the structure.

4.1.2 Effects of High Pressure on DNA/RNA Stability

Even though the number of studies on the effects of high pressure on the physical stability of nucleic acids is comparatively small compared to that done on proteins, several important studies have been made. It has been observed that, at pressures of up to 200 MPa, pressure has very little effect on the structure of a B-DNA helix [34]. However, at pressures of 1 GPa, it was observed by Kryzyzaniak *et al.* that a poly[d(G-C)] DNA sequence, which is in the B-form at atmospheric pressure, undergoes a reversible transition to the Z-form [71]. Changes in the catalytic activity of RNA have also been observed. Fedoruk-Wyszomirska *et al.* have shown that the hammer head ribozyme is capable of catalytic activity under high-pressure conditions in the absence of Mg^{2+} ions (at atmospheric pressures a metal ion is required in the activation site of this ribozyme for catalytic activity) [31]. This result shows that even the small effect of pressure on a nucleic acid structure can still lead to significant consequences in function.

A small amount of work has been performed on the effects of high-pressure conditions on the thermal stability of DNA. Studies by Wu and Macgregor [72, 73] have shown that pressures of 200 MPa can increase the melting point of long synthetic poly[d(A-T)] and poly[d(G-C)] sequences by up to 10 K. It was also observed that the salt concentration plays an important role, with higher salt concentrations producing larger effects. These results are insightful but are limited to the DNA sequences involved. Both of these sequences are long and consist of a simple sequence of alternating bases.

There is still an absence of data on the thermal stability of short DNA sequences of somewhat random, mixed base pairs. These short but more complex sequences

may possibly produce results different to those already observed. With the possibility that high-pressures may offset the negative effects of temperature in the physical stability of RNA and DNA, a more thorough study should be undertaken.

The following is a comprehensive study on the effects of high pressure on the thermo-stability of a short self-complementary DNA hexamer 5′-CAT ATG-3′. ^{1}H NMR spectroscopy was used to study the effects of high pressures on the melting points of **each** residue in this sequence to determine the average melting point and to examine whether these effects are different for different base pairs.

4.2 Materials and Methods

4.2.1 Sample Preparation

The DNA hexamer 5′-CAT ATG-3′ was purchased from Integrated DNA Technologies. This DNA, whilst cleaned, still contained impurities from the manufacturing process (particularly the protecting group, benzamide), which were strongly visible in the ^{1}H NMR spectrum. To remove these impurities, the DNA was first purified by precipitation in cold ethanol.

A pH 7.00 0.20 M phosphate buffer solution was first prepared by adding 0.20 M disodium hydrogen phosphate stock solution to 0.20 M monosodium dihydrogen phosphate stock solution (Table 4.1). The disodium hydrogen phosphate was added in the presence of a pH meter to ensure the correct pH of 7.00.

The DNA was dissolved in 100 µL of milli-Q water and combined with 150 µL of the pH 7.00 0.20 M phosphate buffer, 30 µL of 5 mM TMSP-d4 stock solution (NMR reference) and 60 µL of D_2O (MagniSolv). The solution was made up to 600 µL using milli-Q water, giving final concentrations of: 0.23 mM (confirmed with a nanodrop spectrophotometer) for 5′-CAT ATG-3′, 25 mM phosphate buffer and 0.25 mM TMSP-d4 in H_2O/D_2O (90:10 v/v).

Stock Solution	Mass used /g	Volume used /mL	Concentration /M	Molar Mass /g mol^{-1}	Supplier
NaH_2PO_4. $2H_2O$	1.5601	50.0	0.100	156.01	Riedel-de Haën
Na_2HPO_4. $2H_2O$	1.7799	50.0	0.100	177.99	Riedel-de Haën
TMSP-d4	0.0073	10.00	0.0050	156.26	Merck

Table 4.1: Table of Stock Solutions for Hexamer Melting Experiments. Each entry shows concentration, mass of compound used, volume made (in milli-Q water) and supplier.

It should be noted that there was no attempt to correct for any pressure-induced changes of the pH of the buffer. It has been established that the melting point of DNA is independent of pH over a range of pH 6.5 to 8.3 [74] with only minimal effects over the range of pH 5 to 9 [75]. pH values present at pressures up to 150 MPa are within the pH 6.5 to 8.3 range with pH values at pressures of 200 MPa and 250 MPa only just outside this range (pH 6.29 and pH 6.14 respectively).

4.2.2 DNA Analysis

Circular dichroism (CD) melting experiments were performed on an Applied Photophysics Chirascan CD spectrometer. A 0.1 mm sample cell was used to reduce the path length through the sample to prevent over saturation without diluting the sample. Melting experiments were performed by recording multiple CD spectra over a temperature range of 278.2 K to 353.2 K in 5 K steps. A temperature equilibration time of 5 minutes was used.

NMR analysis of 5′-CAT ATG-3′ was performed with a Bruker Avance 500 MHz NMR spectrometer. Excitation sculpting [76] set at the water offset frequency was used to suppress the dominant water peak in the spectrum.

2D spectra for sequence determination were recorded at 277.2 K using a QXI 5 mm ^{1}H-detection inverse probe with a z-field gradient coil. For these initial

experiments, used for ^{1}H NMR spectrum assignments, the solution was placed in a standard 5 mm glass NMR tube to increase the sample volume and, therefore, give a better signal-to-noise ratio.

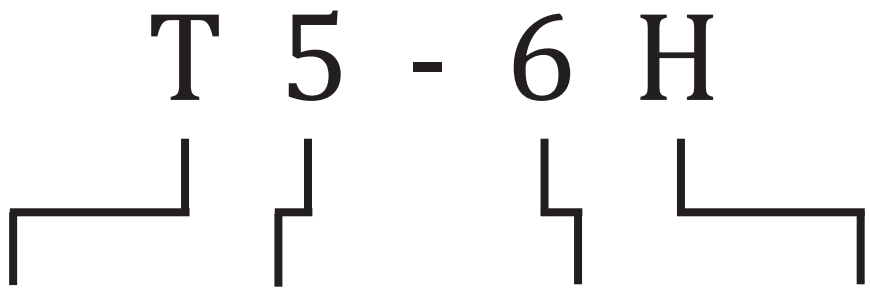
DNA melting experiments were performed by recording multiple NMR spectra at different temperatures using a TXI 5 mm ^{1}H-detection inverse probe with a z-field gradient coil. For the determination of melting points under high-pressure, the DNA solution was transferred to the high-pressure NMR cell. NMR spectra were recorded with 3 K temperature steps from 277.2 K to 334.2 K with an equilibration time of 5 minutes. All 1D spectra were recorded with 512 scans, relaxation time of 2 s, spectral width 20.7 ppm and 64000 points. All 2D spectra were recorded with 512 scans, relaxation time of 2 s, a recycle delay of 2 s and 4096 x 512 points. The COSY spectral width was 18.9 x 19.0 ppm while the NOESY spectral width was 18.9 x 10.2 ppm with a mixing time of 0.5 s.

4.2.3 Spectral Assignment

Examples of the labelling nomenclature used during the course of this chapter are given below. For ^{1}H NMR proton signals, the following format is used:

Indicates that this is for a proton signal	Signal number starting with most down field peak	Base type (A,C,G,T)	Proton number (Figure 4.1)	Proton type (H,CH_3)

Similarly, for assigned sequence positions, the following format is used:



Base type (A,C,G,T)	Sequence position starting at the 5′ end	Proton number (Figure 4.1)	Proton type (H,CH_3)

In some instances the sequence position will be excluded when referring to unassigned protons. Only the proton number and type will be given when referring to generic protons.

In the ^1H NMR spectrum of a DNA strand, the simplest and most distinct signals that can be assigned are those corresponding to the hydrogens on the DNA bases themselves. While most of these signals tend to lie within a specific spectral range (Figure 4.2), it can be near impossible to determine the assignment of two identical bases. For example, a DNA strand with two adenosine bases will exhibit two peaks in the adenosine region but there is no clear indication as to which peak belongs to which base. With the use of the 2D NMR spectroscopy techniques NOESY and COSY (nuclear Overhauser effect spectroscopy and homonuclear correlation spectroscopy), it is possible to identify peaks corresponding to hydrogens A-8H, T-6H, G-8H and C-6H. These signals can then be used to assign the signals for hydrogens A-2H, T-5CH_3 and C-5H.

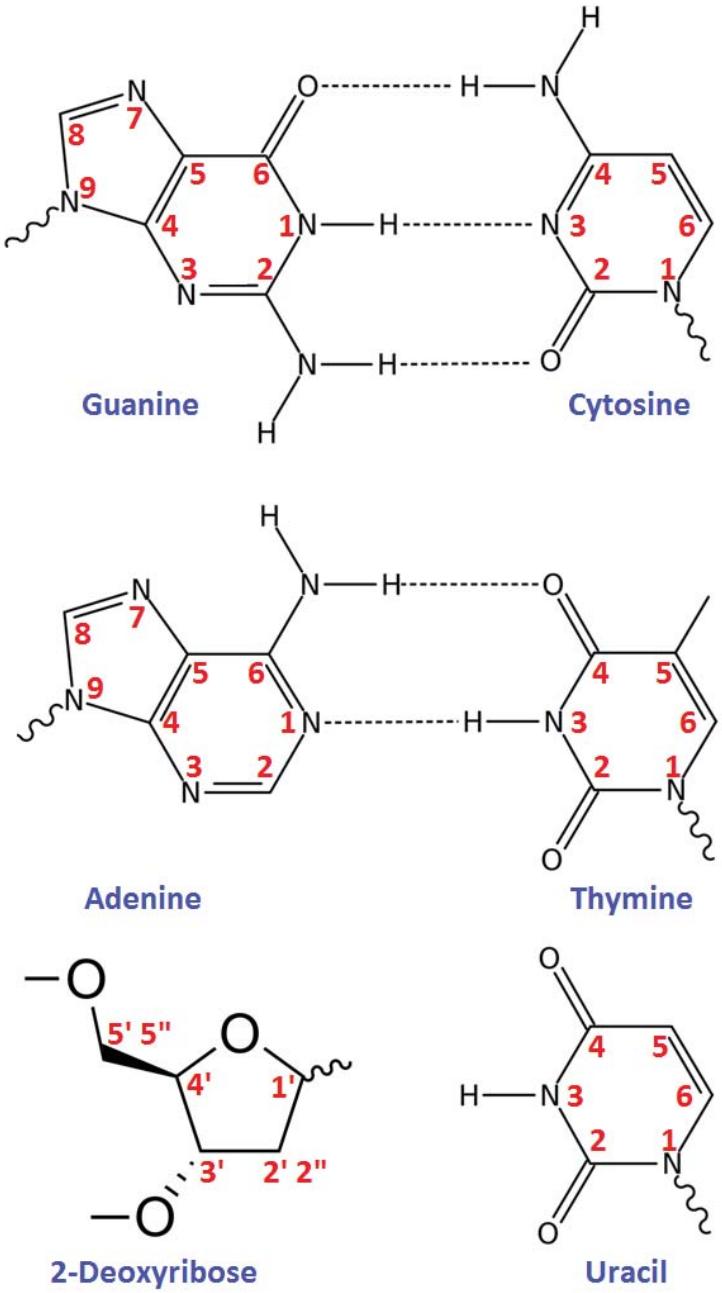


Figure 4.1: Nucleobase Labelling. Proton assignment diagram for all bases including labelling for the deoxyribose sugar. n′ refers to a proton on the 2-deoxyribose group while n″ refers to the second of two hydrogens attached to the same carbon.

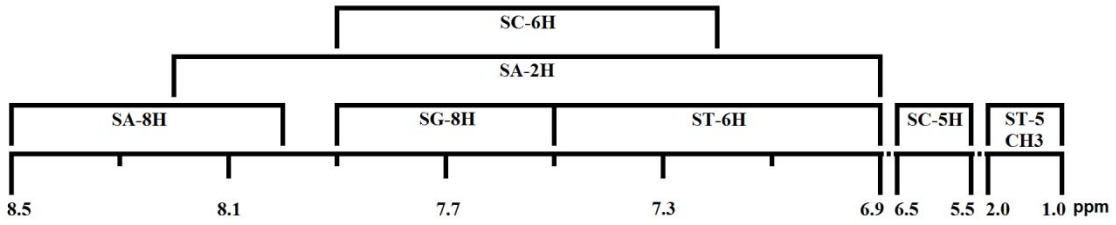


Figure 4.2: Observed Proton Spectral Ranges. The spectral ranges over which specific proton signals occur as observed over the course of this study. Conveniently, C-6H signals can be identified as doubles while A-2H exhibit no NOE's with 2′H/2″H protons in a NOESY spectrum. This allows for the clear identification of the A-8H, G-8H, T-6H and T-5CH_3 signals. These results compare well with results seen by Nielsen *et al.* [77].

To obtain appropriate NMR spectra for spectral assignment of the DNA sequence, the following method was used.

- A brief melting experiment was first carried out to both, obtain an initial indication of the DNA's melting point and, more importantly, to determine a temperature at which there is the greatest degree of spectral separation (lowest number of overlapped signals) in the region of interest (8.6-6.9 ppm). This was found to be 277.2 K.

- Both NOESY and COSY 2D NMR spectra of the DNA strand were recorded at 277.2 K. These spectra were used in the assignment for 5′-CAT-ATG.

Assignments were made using the method outlined by Wüthrich [78]. This method uses ^{1}H – ^{1}H nuclear Overhauser effects (NOE) to establish assignment link pathways. These link pathways are formed by alternating intraresidue connectivities, between bases and their sugars, denoted i, and sequential connectivities, between sugars and the neighbouring nucleobase, denoted s (Figure 4.3).

These ^{1}H – ^{1}H connectivities can be denoted as d_i(6H/8H→1′H) and d_s(1′H→6H/8H). NOE cross peaks formed from these connectivities in a NOESY spectrum can be linked in order of their occurrence. With the nucleobase region (8.6-6.9 ppm) on the *x* axis, this can be seen as a series of connected horizontal lines (i – s connectivities) and vertical lines (s – i connectivities) (Figure 4.4). By following these connecting lines (links) it is possible to establish an assignment link pathway, which can be used to assign non-labile protons in B-DNA. This method can be used to solve the DNA assignment using only a NOESY spectrum;

however, results can be further confirmed with the use of a COSY spectrum, particularly in longer DNA chains. The sequential assignment link pathways shown in Figure 4.4 have limitations as a result of overlapping signals in 2D NMR spectra. In light of this, three additional pieces of information can be used to assist with the spectral assignment.

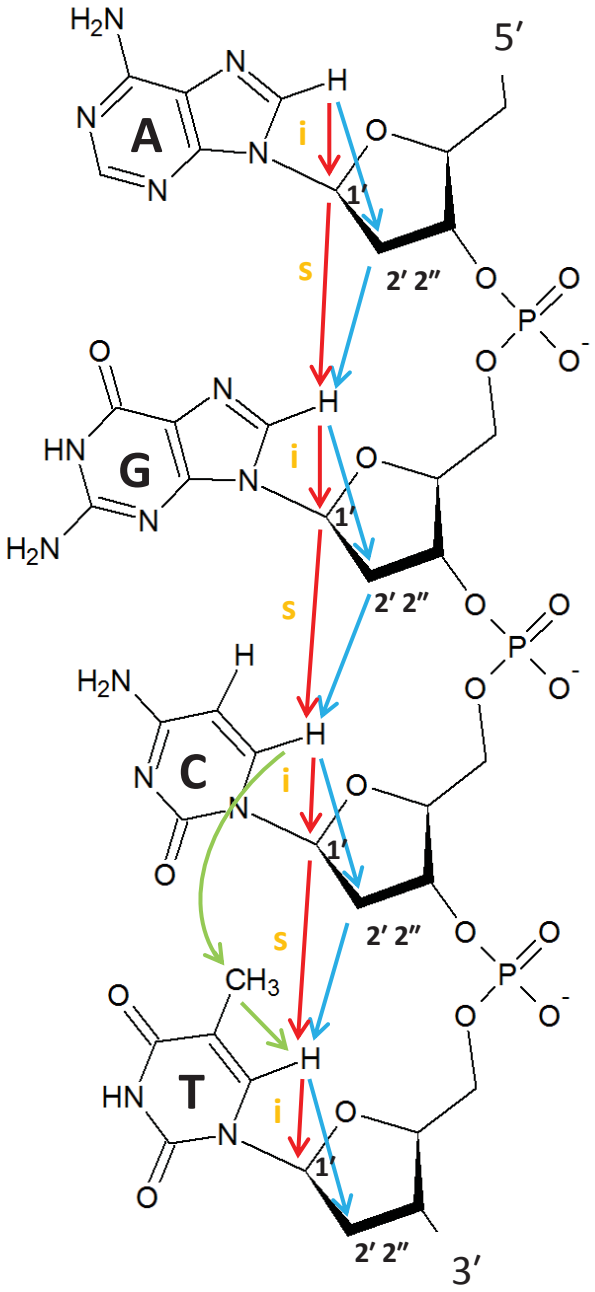


Figure 4.3: Intraresidue (i) and Sequential (s) Connectivities Between Adjacent Nucleotides. The ^{1}H – connectivities d_i(6H/8H→1′H) and d_s(1′H→6H/8H) (red arrows), d_i(6H/8H→2H′)/d_i(6H/8H→2H″) and d_s(2′H→6H/8H)/d_s(2″H→6H/8H) (blue arrows) and d_s(6H/8H→5CH_3) (green arrows) are used for sequential assignments of non-labile protons in B-DNA. Image adapted from Wüthrich [78].

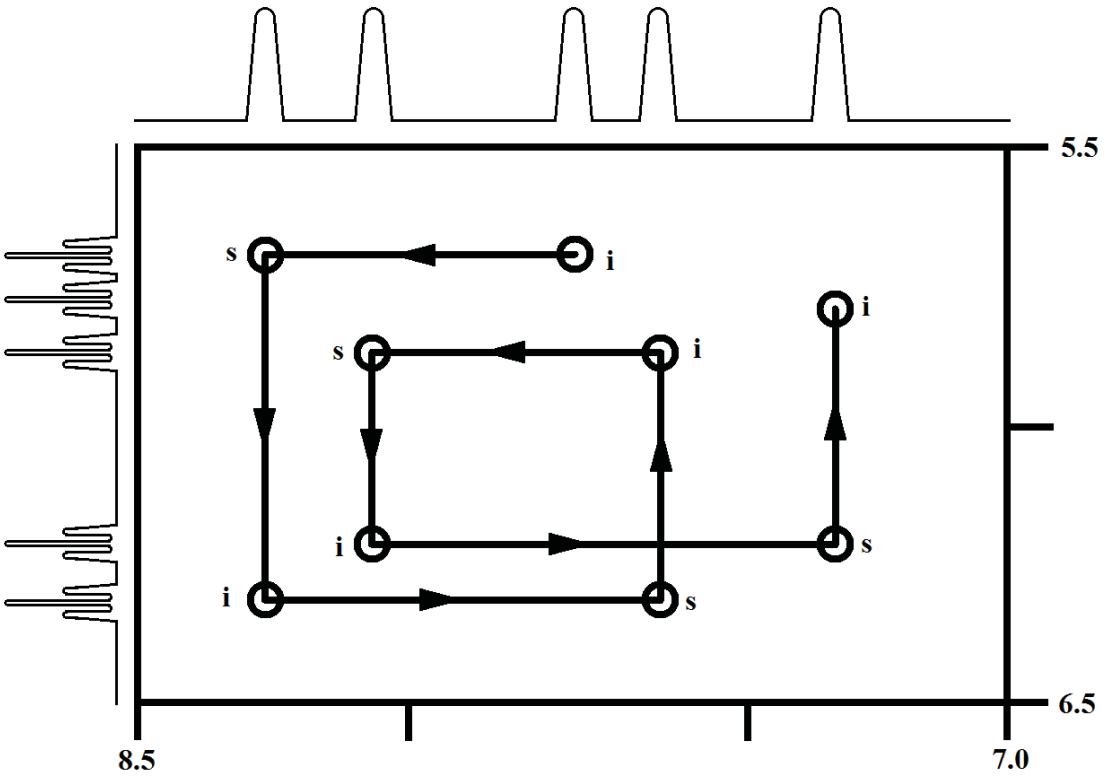


Figure 4.4: Example of Cross Peaks and Their Links. Examples of the NOE cross peaks and their links in the anomeric (6.5 to 5.5 ppm)/base proton (8.5 to 7.0 ppm) region. Horizontal lines indicate i – s links and vertical lines indicate s – i links.

Terminal bases can be identified and used as reference locations for sequential assignments. The 5′-terminal base and the 3′-terminal deoxyribose are unique in exhibiting only intranucleotide NOE's. This results in no vertical links from the 5′-terminal base and only a single horizontal link for the 3′-terminal deoxyribose (Figure 4.5 a).

Links can be observed between the intranucleotide connectivity $d_i(5CH_3 \rightarrow 6H)$ (between the T-6H and the T-$5CH_3$) and the sequential connectivity $d_s(6H/8H \rightarrow 5CH_3)$ (between T-$5CH_3$ and the 6H or 8H of the preceding nucleotides). These are horizontal links (Figure 4.5 b) in the region of [8.6-6.9:1.8-1.3] ppm. These links identify the nucleotide preceding the thymine and can be used to identify non-terminal sequence locations within the DNA strand.

Other non-terminal sequence locations can be identified by matching partially assembled sequences with corresponding segments in the DNA sequence (Figure 4.5 c). This is similar to techniques used for proteins.

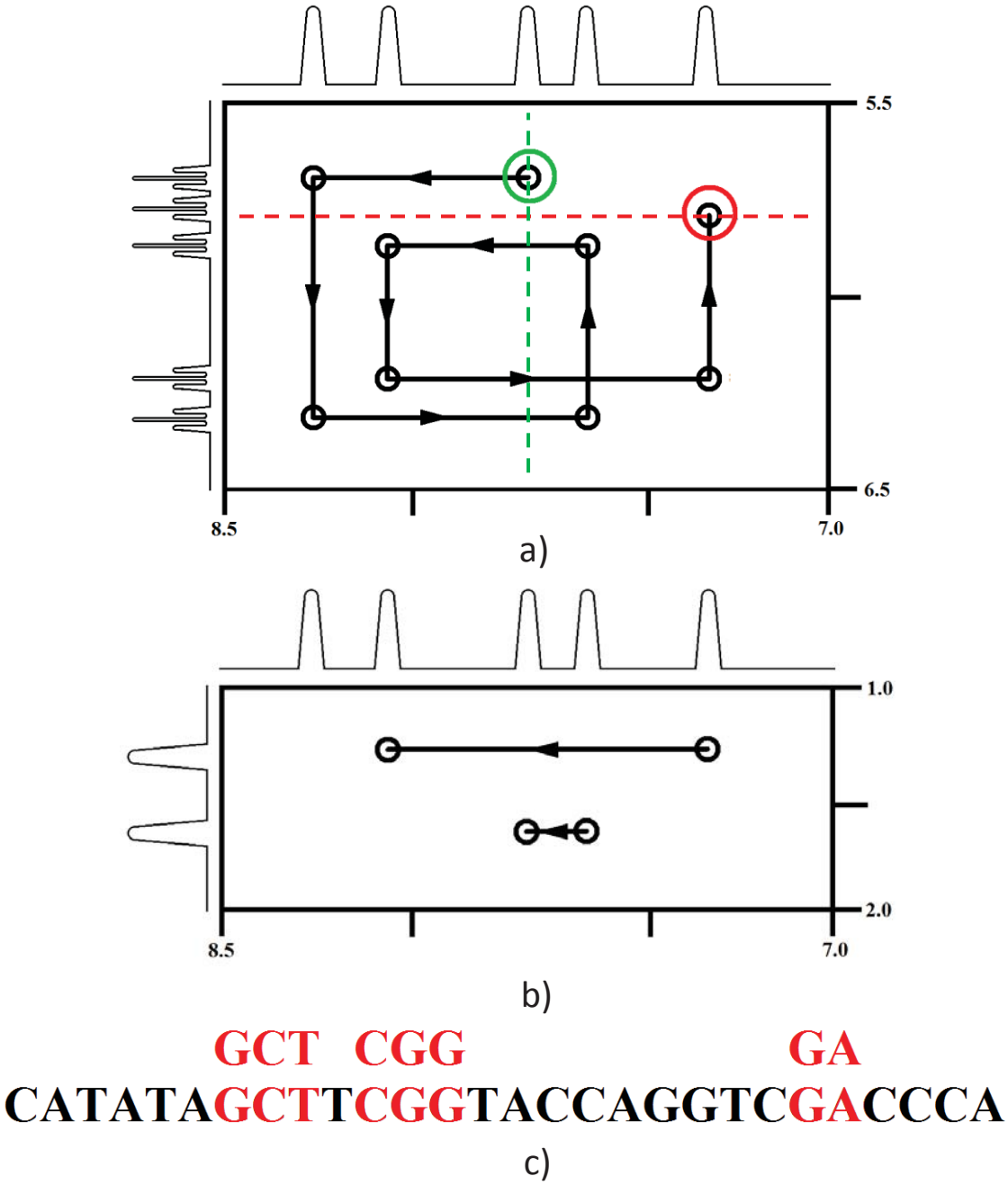


Figure 4.5: Examples of Partial Sequence Determination. a) Identification of terminal bases. The 5′-terminal base (green circle) and the 3′-terminal deoxyribose (red circle) exhibit only intranucleotide NOE's (as indicated by no other signals along the dashed lines). A 3′-terminal cytosine will have a second NOE signal corresponding to the connectivity between the adjacent C-6H, C-5H protons. This signal tends to be stronger than other signals and should not be confused as a 1′H, C-6H NOE. b) NOE links (horizontal lines) between the intranucleotide connectivity $d_i(5CH_3 \rightarrow 6H)$ and the sequential connectivity $d_s(6H/8H \rightarrow 5CH_3)$. c) Partial sequences being matched to the DNA sequence. The remaining signals can then be identified more easily.

The assignment of the ^{1}H NMR spectrum of 5′-CAT ATG-3′ was made using the d_i(6H/8H→1′H) and d_s(1′H→6H/8H) sequential assignments. This was confirmed with d_i(6H/8H→2′H)/d_i(6H/8H→2″H) and d_s(2′H→6H/8H)/d_s(2″H→6/8H) sequential assignments from both NOESY and combined NOESY-COSY spectra.

4.2.4 Spectral Assignment of 5′-CAT ATG-3′

The ^{1}H NMR spectrum of 5′-CAT ATG-3′ can be assigned as follows:

With the use of Figure 4.2 and a NOESY spectrum, the type of base for each signal can be identified (Figure 4.6). This was done by first identifying C-6H and A-2H peaks. Cytosine signals show unique doublets while A-2H peaks exhibit no NOE's with 2′H/2″H protons. Once these were eliminated the remaining signals can be identified. This was done by simply examining the chemical shift of each peak and comparing to the ranges given in Figure 4.2.

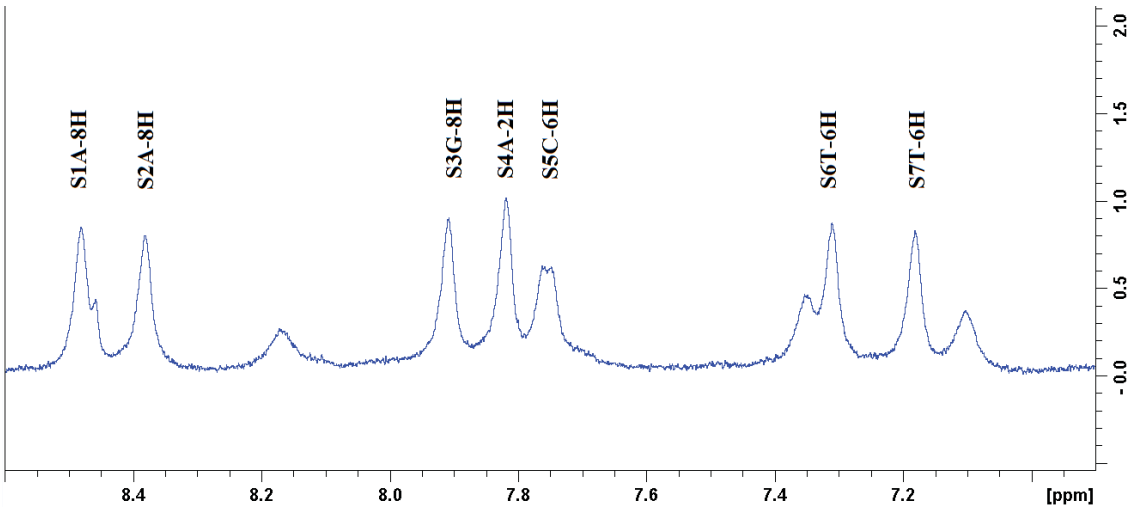


Figure 4.6: ^{1}H NMR Spectrum of 5′-CAT ATG-3′. Protons have been assigned as described above.

The terminal bases were then identified. In this instance this was particularly simple since there is only one of each cytosine and guanine base. Regardless, it is

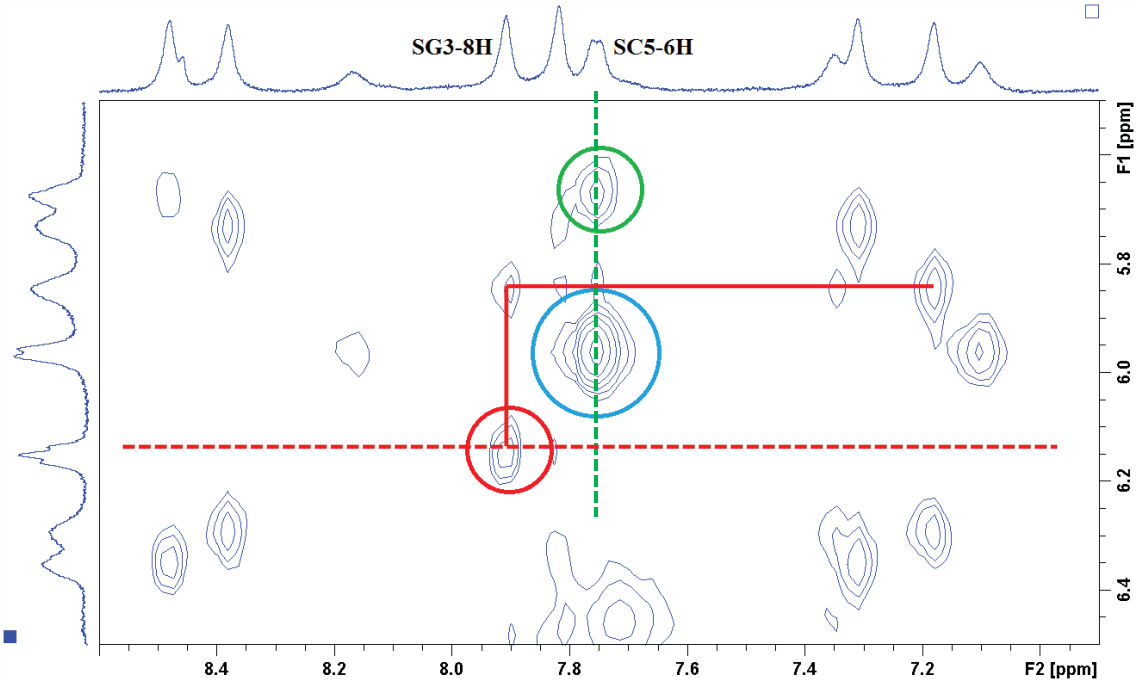


Figure 4.7: Terminal Bases for 5′-CAT ATG-3′. The 5′-terminal cytosine base (green circle) and the 3′-terminal guanine deoxyribose (red circle). Note that there are no additional cross peaks along the dashed lines (the large cross peak on the vertical line (blue circle) is from the connectivity between C-6H, C-5H protons

clear for each of these that there are only intranucleotide interactions with the 5′-terminal base and the 3′-terminal deoxyribose (Figure 4.7).

Two links from the CH_3 groups on the thymine residues can be clearly identified between S1A-8H→CH_3 – CH_3→S6T-6H and S2A-8H→CH_3 – CH_3→S7T-6H (Figure 4.8). This gives the two expected AT sequences but does not distinguish between the two.

Using the information above, an assignment link pathway using d_i(6H/8H→1′H) and d_s(1H′→6H/8H) connectivities can be determined (Figure 4.9) and the DNA sequence can be assigned as follows (Table 4.2).

This result is confirmed with d_i(6H/8H→2′H)/d_i(6H/8H→2″H) and d_s(2′H→6H/8H)/d_s(2″H→6H/8H) sequential link pathways with both NOESY and combined NOESY-COSY connectivity diagrams as shown in Figure 4.10 and Figure 4.11.

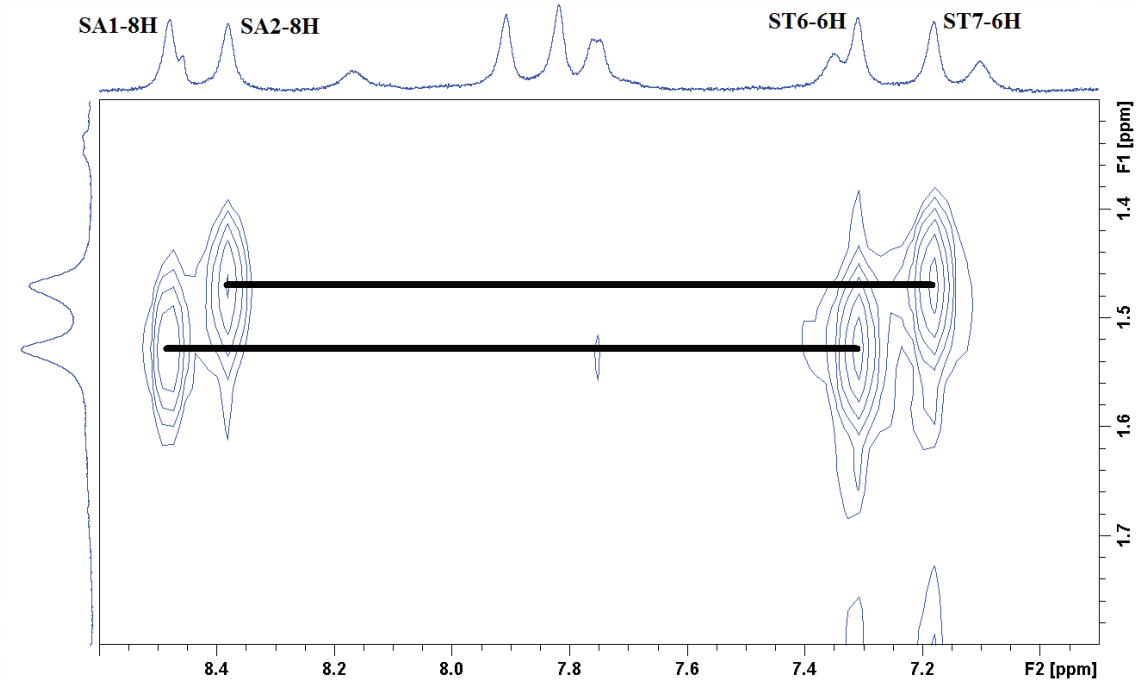


Figure 4.8: S1A-8H→CH_3 – CH_3→S6T-6H and S2A-8H→CH_3 – CH_3→S7T-6H Links. Links are indicated by horizontal lines.

Base and assigned sequence number	C1	A2	T3	A4	T5	G6
Signal	S5C-6H	S1A-8H	S6T-6H	S2A-8H	S7T-6H	S3G-8H
Frequency at 277.2 K	7.752	8.476	7.303	8.375	7.180	7.908

Table 4.2: Sequencing Assignment for 5′-CAT ATG-3′.

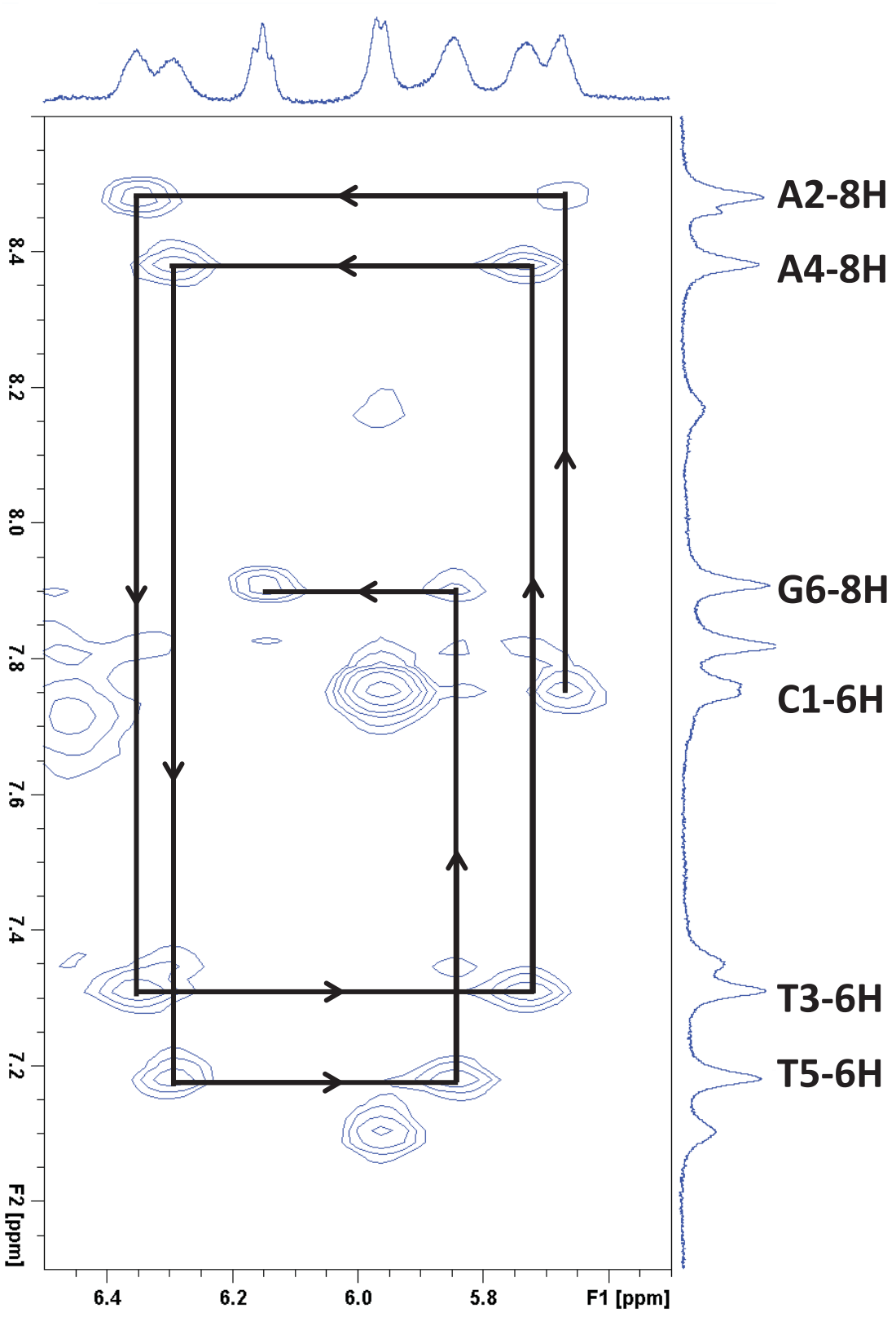


Figure 4.9: NOESY Assignment Link Pathway for 5′-CAT ATG-3′ with d_i(6H/8H→1′H) and d_s(1′H→6H/8H) Connectivities.

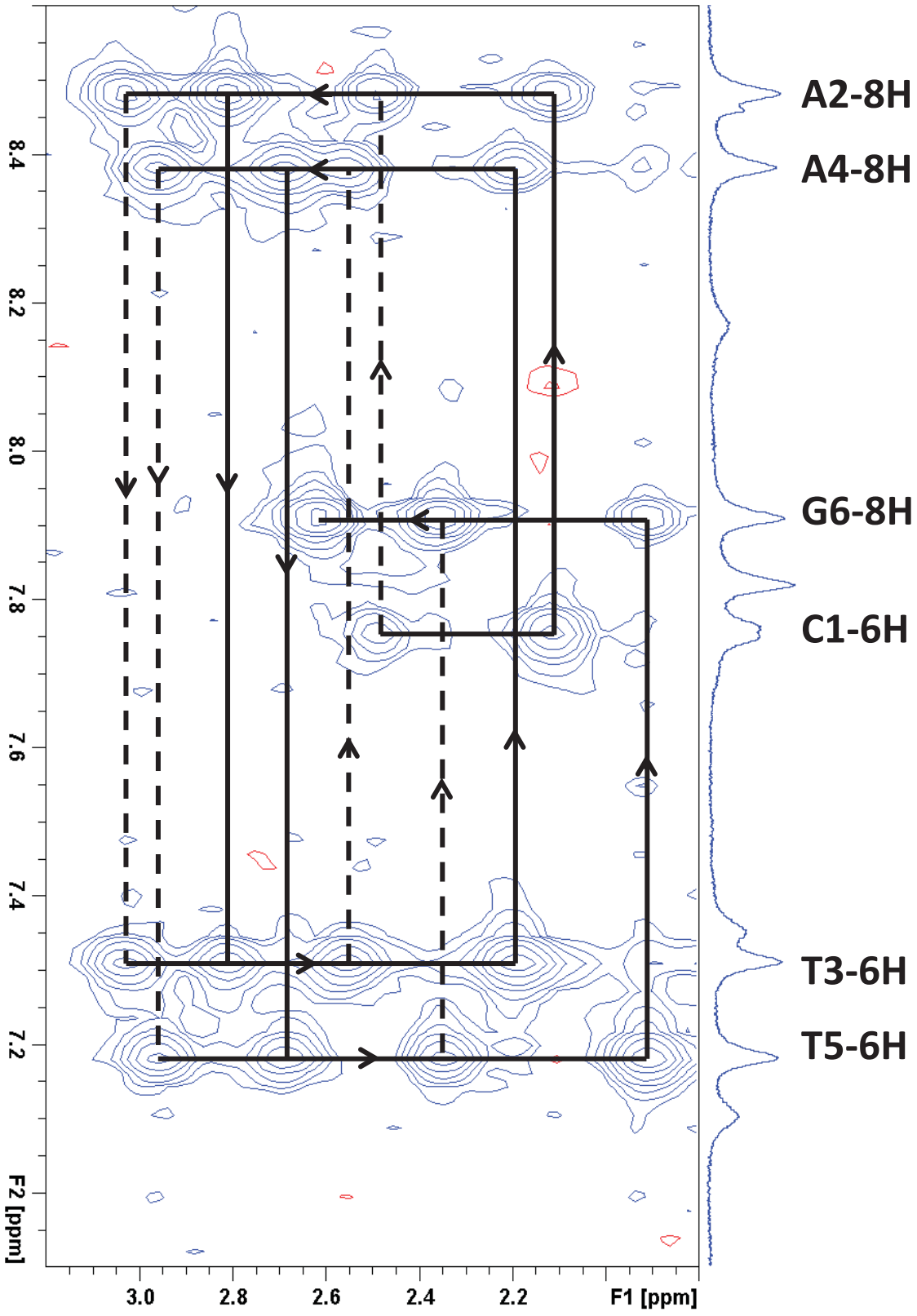


Figure 4.10: NOESY Assignment Link Pathway for 5′-CAT ATG-3′ with d_i(6H/8H→2′H)/ d_i(6H/8H→2″H) and d_s(2′H→6H/8H)/d_s(2H″→6H/8H) Connectivities. Solid horizontal lines indicate 2′H i-s links and horizontal dashed lines represent 2″H i-s links (as seen side on).

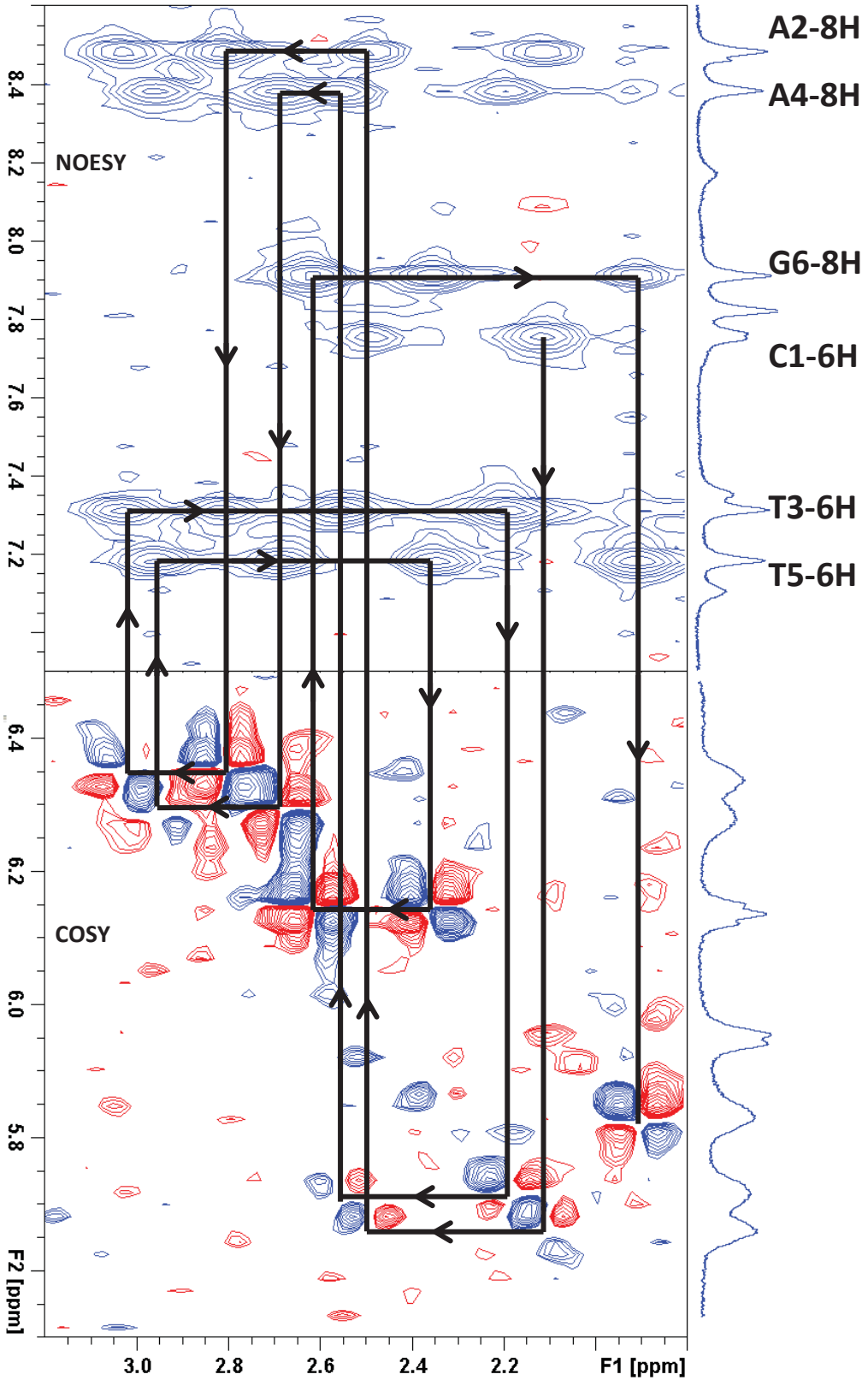


Figure 4.11: Combined NOESY-COSY Assignment Link Pathway for 5′-CAT ATG-3′ with d_i(6H/8H→2′H) and d_s(2″H→6H/8H) Connectivities. The signals forming the COSY spectrum are each comprised of two symmetrical cross-peaks (one positive, blue, and one negative, red), each along the one of the diagonals of the spectra. These cross peaks are a result of magnetisation transfer between two coupled protons. Where these diagonals cross is the centre of the overall signal. Note: some of these cross-peaks overlap or are not fully visible at this resolution.

4.2.5 Melting Experiments

After successfully assigning the ^{1}H NMR spectrum, high pressure melting experiments were performed. These experiments involved heating of the sample in stepped temperature increments while recording a ^{1}H NMR spectrum at each step. This process was repeated at pressures of 0.1, 100, 150, 200 and 240 MPa using the high-pressure cell. These NMR spectra were then used to determine the melting point for 5′-CAT ATG-3′ under different pressure conditions. Melting experiments done at 0.1, 100, 200 and 240 MPa had duplicate runs while experiments done at 150 MPa were run in triplicate.

4.2.6 Circular Dichroism Experiments

A circular dichroism melting experiment was performed in order to obtain an accurate melting point to confirm the results obtained using the ^{1}H NMR spectroscopy. A 50 µL drop of the DNA solution was placed in a 0.1 mm quartz cell. The CD spectrum of the solution was recorded at stepped temperature increments at 0.1 MPa. These spectra were then used to determine the melting point of the DNA to compare to the results from the NMR spectroscopy.

4.3 Results

4.3.1 Data Gathering and Processing

For each spectrum obtained at 0.1 MPa, the chemical shifts for the signals corresponding to protons C-6H, A-8H, A-2H, T-6H, T-5CH_3 and G-8H were recorded through the melting of the DNA. During this process, the peaks broadened and moved considerably (Figure 4.12). To avoid confusion, each peak had to be carefully monitored through the melting process so as to obtain the correct curve.

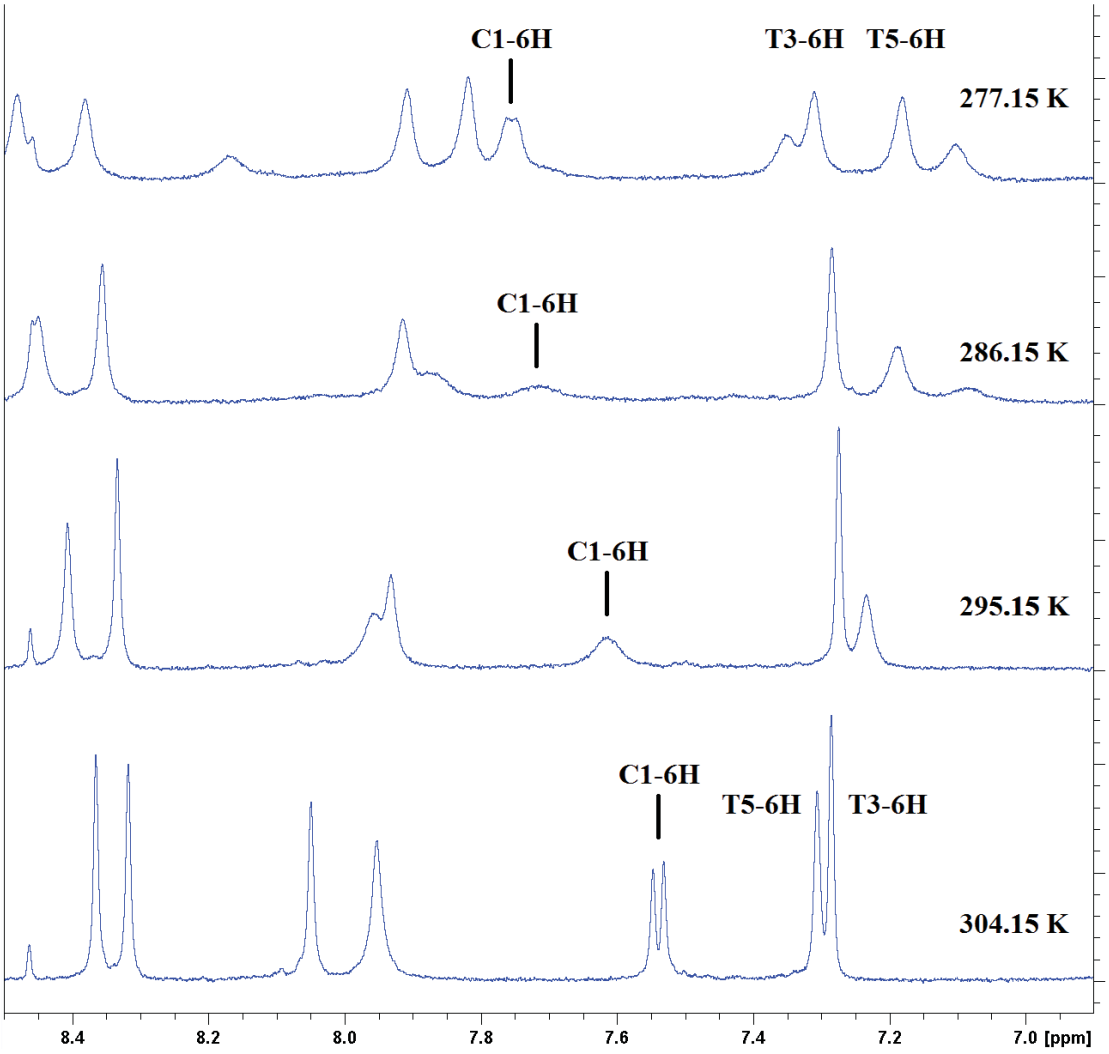


Figure 4.12: Peak Broadening and Shifting. An example of the broadening of H1-NMR peaks during the melting process of 5′-CAT ATG-3′. The C1-6H signal starts as a sharp peak in the duplex form at 277.2 K before it broadens until almost flat at 286.2 K. It then proceeds to sharpen again in the final monomer form at 304.2 K. It can also be seen that the T3-6H and T5-6H peaks have crossed over each other.

This resulted in a plot of the chemical shift of each proton as a function of temperature). These data (melting curves) were then normalised. This was then repeated for each pressure condition. Melting curves obtained for each residue studied at each pressure condition are presented in Figure 4.13 .

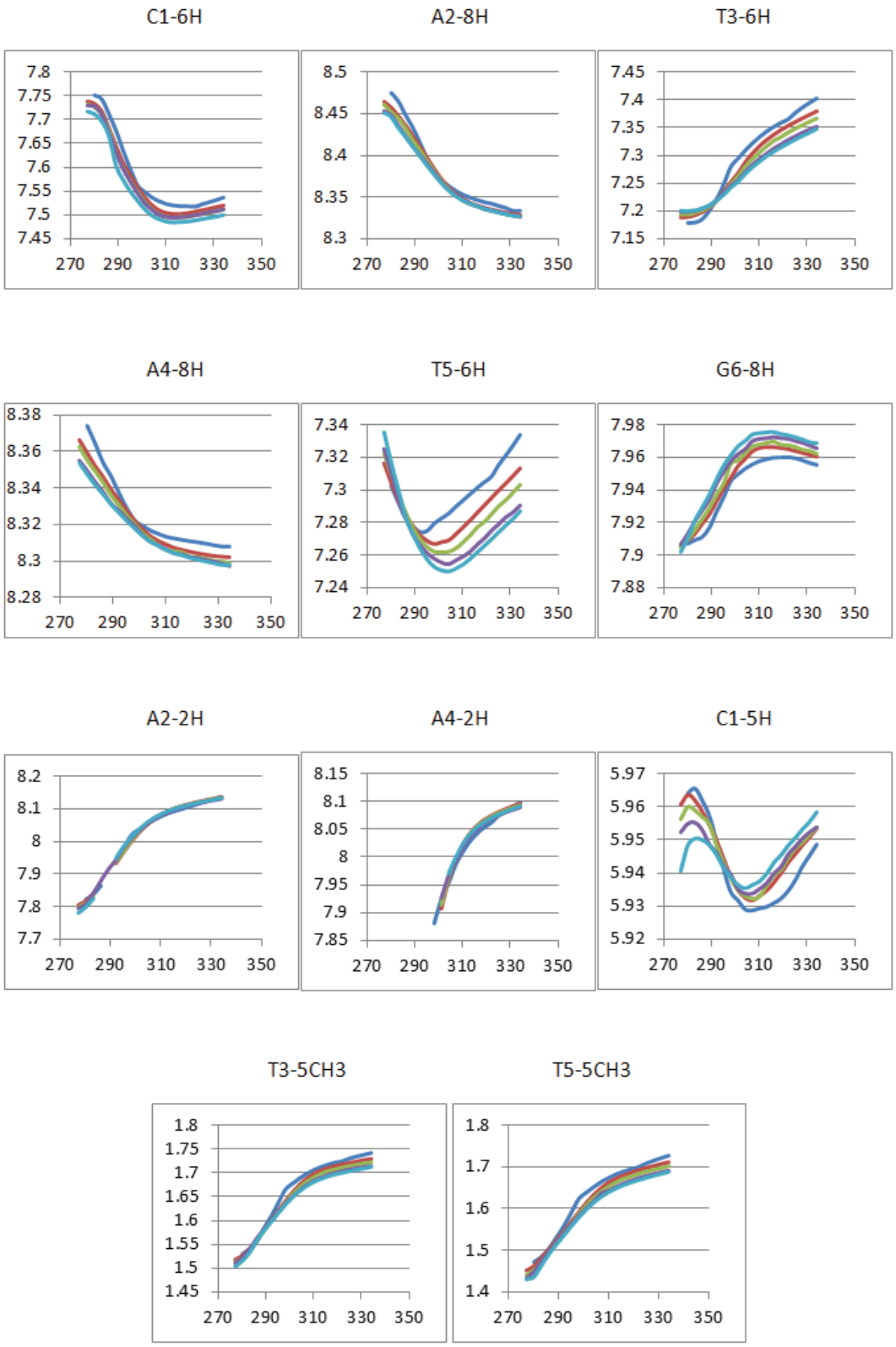


Figure 4.13: Residue Melting Curves for the DNA Hexamer. 0.1 MPa – blue, 100 MPa – red, 150 MPa – green, 200 MPa – purple, 240 MPa – light blue. Melting curves are presented as the chemical shift (/ppm) as a function of temperature (/K).

4.3.2 Data Fitting

After data processing, the resulting melting curves were fitted to a sigmoidal function to obtain the melting point.

$$y(x) = \frac{1}{1 + e^{-x}} \tag{19}$$

Equation (19) was adapted to suit the fitting of the resulting melting curves by inserting the appropriate variables *a, b*, *c* and *d* to give:

$$y(x) = \frac{a}{1 + e^{b(x-c)}} + d \tag{20}$$

Equation (20) has been constructed in such a way as to allow the variable *c* to be equal to the melting point of the fitted melting curve.

To account for the small overall gradient present in the melting curve an extra term, *xe*, was also added.

$$y(x) = \frac{a}{1 + e^{b(x-c)}} + xe + d \tag{21}$$

Initial attempts to fit the melting curves using equation (21) were only partially successful. While some curves fitted well others were unable to be fitted (Figure 4.14).

It is clear that, while the sigmoidal function in equation (21) has an added gradient function to account for the steady change in chemical shift of the proton signals, many of the melting curves possess different gradients before and after the melting curve. It was proposed that a second sigmoidal function be added to replace the gradient function in equation (21) which would, in effect, apply two separate gradient functions before and after the central sigmoidal curve. This leads to:

$$y(x) = \frac{a}{1 + e^{b(x-c)}} + x\left(\frac{x(f-e)}{1 + e^{b(x-c)}} + f\right) + d \tag{22}$$

where *e* and *f* are the gradients of the melting curve at the beginning and the end of the melting curve, respectively. This can be simplified to:

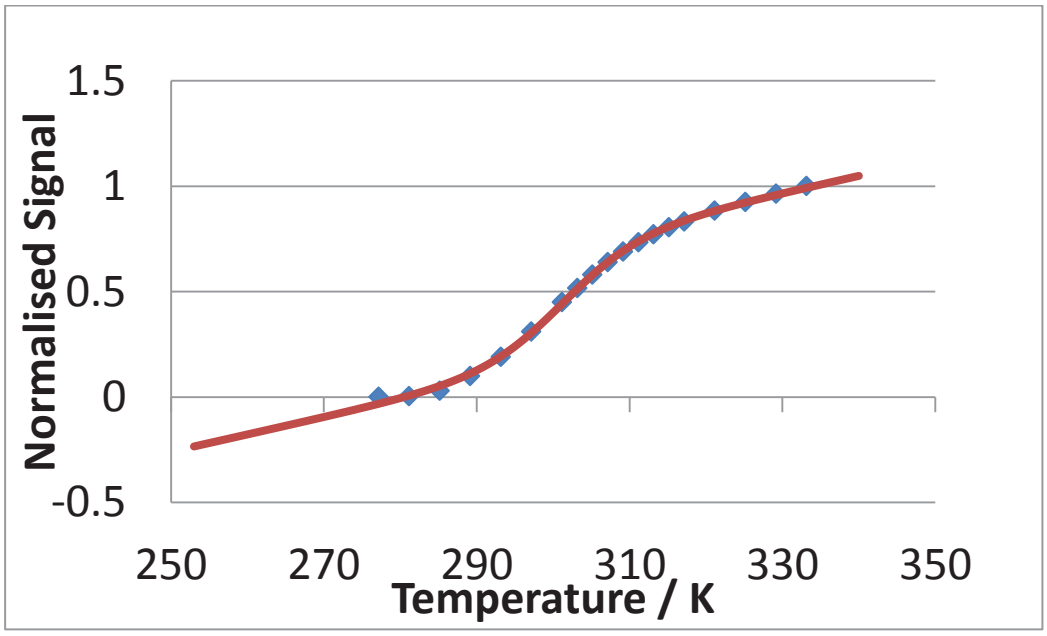


Figure 4.14: An Incorrectly Fitted Curve. The melting curve (blue) of T3-6H has been fitted with equation (21) (red). The gradient of the fitted line below 280 K does not follow the trend of the line.

$$y(x) = \frac{a - x(f - e)}{1 + e^{b(x-c)}} + xf + d \tag{23}$$

Equation (23) resulted in a better fit of the acquired melting curves (Figure 4.15).

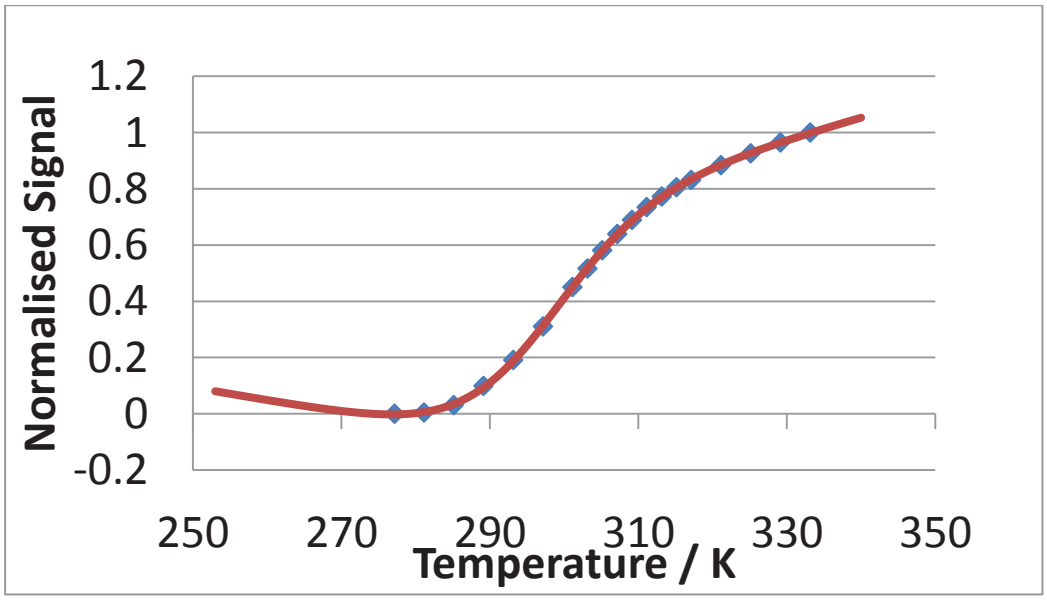


Figure 4.15: Improved Fit. The melting curve (blue) of T3-6H has been fitted with equation (23) (red). The gradient of the fitted line below 280 K now has a more suitable gradient.

Using equation (23) the melting curves for each proton signal at each pressure condition were fitted using Gnuplot [79] and the melting points determined. Fitting parameters are available in the digital Appendices.

For some curves it was not possible to successfully determine the melting point from the melting curve by fitting of equation (23). This was a result of incomplete melting curves caused by the inability of the high-pressure cell to go below 277.2 K. In these instances the melting curve was fitted empirically to a polynomial function. The point of inflection in the middle of the curve was then determined by taking the double derivative of this function, giving the melting point.

The resulting plot of average melting point as a function of pressure for each proton signal is displayed below in Figure 4.16.

During the melting process, it was not possible to follow the ^{1}H NMR signal corresponding to A4-2H. In the monomer state, this signal was clearly visible but could not be accounted for in the duplex state (either too broad or has shifted significantly). As a result of this, it was only possible to obtain half of the melting curve (Figure 4.13) and therefore the melting points for A4-2H can only be considered as approximations. These melting points have been displayed as a dashed line in Figure 4.16 but have not been used in any calculations.

As a result of the averaging of the melting points acquired for each pressure condition, there is an approximate error of 1.0 K for each melting point. Even considering this, there is still a clear decrease in the melting point as a function of pressure.

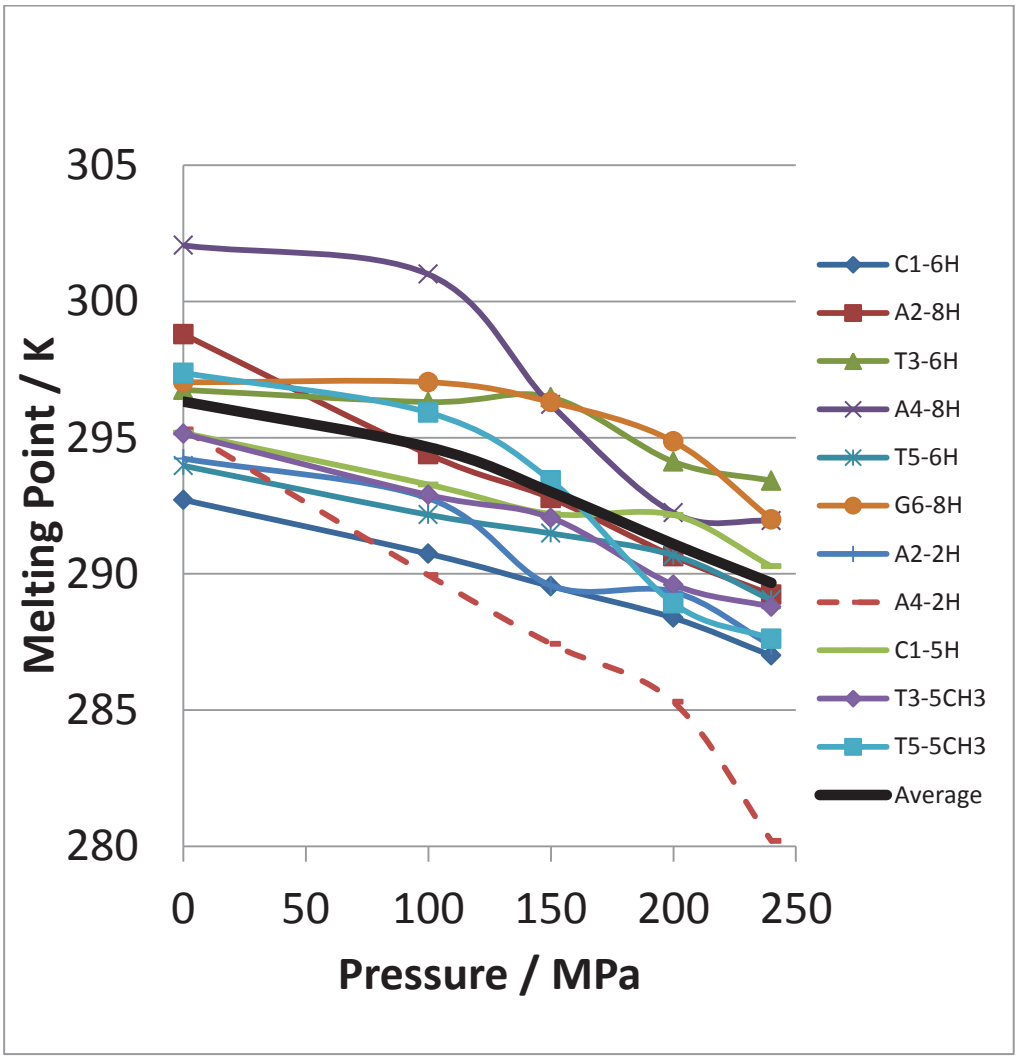


Figure 4.16: Melting Point versus Pressure for 5′-CAT ATG-3′. The curves for each base hydrogen are displayed along with the average melting point (bold black line).

4.3.3 Circular Dichroism Results

Figure 4.17 shows the CD spectra obtained for the melting of the DNA solution. The signal intensities at 208, 270 and 290 nm were recorded, smoothed over 6 nm, normalised and plotted as a function of temperature (Figure 4.18). The melting curve for 250 nm is clearly different from the others. The linear nature of this curve implies that it is caused by a different, noncooperative phenomenon rather than simple, cooperative melting. This is suspected to be caused by the terminal bases whose effects will be more noticeable due to the short sequence length. This curve has been excluded from the calculation of the overall melting point. The

formation of bubbles has caused the non-perfect isosbestic points in Figure 4.17. The melting point of the DNA can determined from the average temperature corresponding to an intensity of 0.5. The average melting point determined from the CD measurements is 297 ± 2 K. This compares quite well with the average result obtained via NMR spectroscopy at 0.1 MPa of 296 ± 2 K.

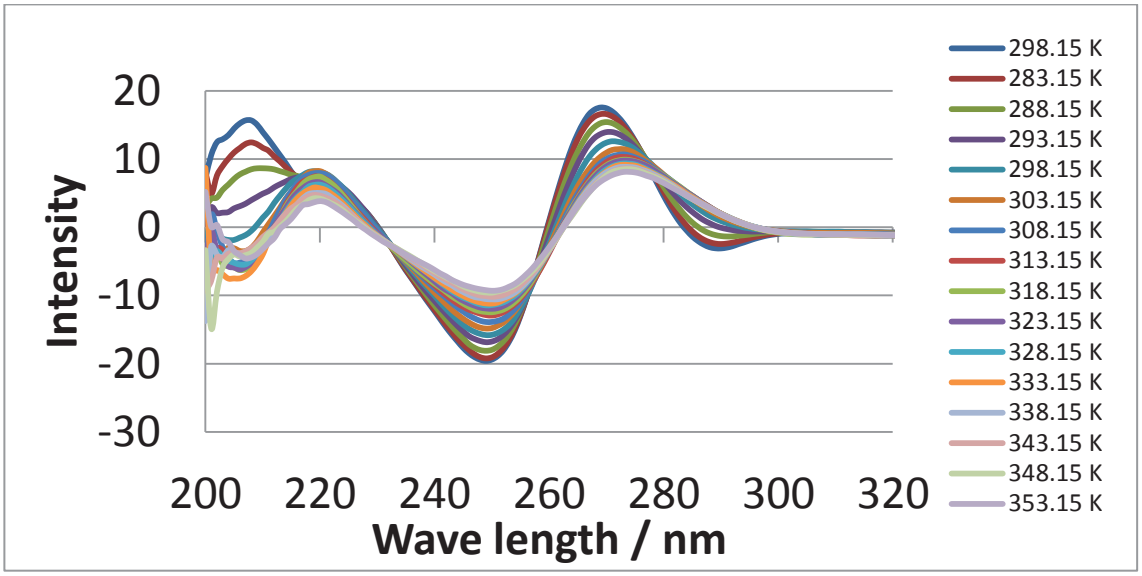


Figure 4.17: CD Spectra During DNA Melting. The intensities at 208, 270 and 290 nm can be used to determine the melting point of the DNA.

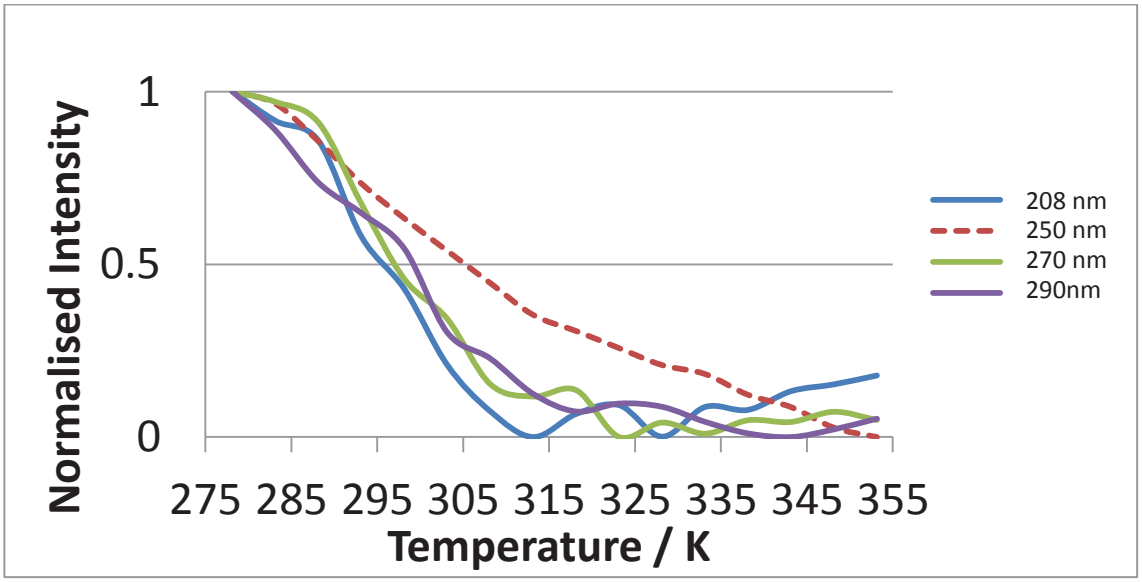


Figure 4.18: CD Melting Curves. Melting curves obtained via CD spectroscopy at wavelengths of 208, 250, 270 and 290 nm. The melting point can be determined by the temperature corresponding to a normalised intensity of 0.5.

4.4 Discussion

The effect of high temperatures on the physical stability of DNA has been well characterised. While most DNA sequences which exceed 100 base pairs are physically stable to temperatures above 80 °C, short sequences are less stable and are therefore not likely to exist in their duplex form in hot or even warm conditions. To better understand the effects of the physical environment on the physical stability of a small DNA sequence, studies were undertaken to determine the effects of high pressures on the physical stability of 5′-CAT ATG-3′.

Figure 4.16 displays the results of melting point versus pressure for each base proton. This figure indicates a clear decrease in the melting point of the DNA sequence as a result of the increased pressure as monitored by the chemical shift of each proton signal. Over the pressure range of 0.1 MPa to 240 MPa, the average melting point decreases from 296.2 K to 288.8 K, a decrease of 7.4 K. This result is a significant change well above the experimental error in the measurements.

The melting temperature at 0.1 MPa has been confirmed by CD spectroscopy to be 297 ± 2 K. This compares well to the result obtained via NMR spectroscopy of 296 ± 2 K.

The decrease in the melting point is not what was initially expected and is in fact the opposite to the results obtained for large simple DNA sequences by Wu and Macgregor [72, 73]. It may be expected that the increase in pressure will force the nucleobases closer together, increasing the strength of the hydrogen bonds and therefore increasing the melting temperature. However, this decrease in the melting point can be explained if the nature of the DNA sequence is considered.

For the self-complementary DNA sequence, 5′-CAT ATG-3′, the only two C-G base pairs are located at the ends of the sequence. While these pairs have an inherently stronger bond strength due to their three hydrogen bonds, their location at the ends of the sequence results in a lack of stabilising effects from other nucleotides, reducing their bond strength, making the base pair more susceptible to opening. This will reduce the effect of steric strain on the terminal G-C pairs caused by

imperfect stacking of the bases [80] and in turn, reduce the susceptibility of the melting point to pressure for these bases. This is supported by the lower susceptibility of pressure on the melting point of the C, G protons evident in the T_m versus pressure curves. The average reduction in the melting points of the C,G protons between 0.1 and 240 MPa is 3.9 K compared to 7.0 K for the centralised A and T protons.

With the lack of stability in the terminal C-G pairs, the small number of more weakly bound central A-T pairs will be more susceptible to steric forces induced by the small compression which results from the increase in pressure [34]. It would appear that these steric forces outweigh the increase in the hydrogen bond strength caused by the increased proximity of the bases, decreasing the overall stability of the central ATAT region and, therefore, of the duplex.

It should also be noted that the results obtained by Wu and Macgregor are in the presence of NaCl (20 mM to 200 mM) whereas the solutions used in this study contained no salt. This addition of salt could have provided additional stabilisation effects.

This decrease in the stability of 5′-CAT ATG-3′, while significant over the pressure range studied, is relatively insignificant when compared to the pressures of oceanic environments. At pressures of 100 MPa (maximum oceanic pressures are 110 MPa), the decrease in the melting point of this DNA sequence is approximately 2 K. This is quite a small change considering the myriad of other environmental conditions which might affect the physical stability of DNA.

The physical stability of DNA at high temperatures is of obvious concern for the consideration of any origin-of-life theory. While DNA tends to be physically stable over a wide range of temperatures from cold to warm, at temperatures above 100°C, this stability is severely diminished. This effect reduces the likelihood of self-replicating RNA-based life forms developing at high pressure and high temperature, in the direct presence of deep sea hydrothermal vents.

Similar to chemical stability, it had been considered that the effects of high temperatures on the physical stability of DNA may be offset by high pressure conditions. The results obtained here show that, rather than increasing the

physical stability, an increase in pressure decreases the physical stability of DNA. Over large pressure ranges this decrease in stability is significant; however, it is still insignificant at more common oceanic pressures. Therefore, it can be concluded that these results, rather than favouring, disfavour, albeit by a small degree, a high-pressure origin of life.

4.5 Conclusions

The results obtained on the effects of high pressures on the thermal stability of the short DNA hexamer 5′-CAT ATG-3′ have shown that this DNA sequence is less thermally stable at higher pressures. A decrease in the average melting point of 7.4 K has been observed over the pressure range of 0.1 MPa to 240 MPa with the central ATAT section having a larger shift in their melting points than that of the terminal G-C pairings. However, these changes in thermal stability are relatively minor at oceanic pressures. The decrease in stability was unexpected as previous work has shown that the thermal stability of DNA to increase with pressure. This decrease in the melting point has been attributed to the small number of weak A-T bonds being overwhelmed by other forces.

4.6 Future Work

These first results studying the effects of pressure on the physical stability of a DNA hexamer are only the beginning. There are still many experiments which can be performed examining longer sequences, the effect of G-C content, RNA and the effects of salts. All these different conditions will need to be comprehensively studied if we are to fully understand how pressure affects the thermal stability of DNA.

4.6.1 Longer Sequences

The results obtained from these studies have been limited by the inherent instability of self-association of the short hexamer sequence. It has been considered that more information on the behaviour of the melting point with pressure may be obtained from a longer DNA sequence. Thus the melting point experiments may be repeated with 8-mer or 12-mer sequences.

4.6.2 G-C content

There is a strong possibility that the way in which the melting point of a DNA sequence is affected by pressure could be linked to the sequence itself. Differences in the sequence, or at least the G-C content of the sequence, could alter its melting behaviour.

4.6.3 RNA

Repeating these studies using RNA instead of DNA would be extremely relevant. As described in **Chapter 1**, it is strongly accepted that RNA preceded DNA. Whilst these initial studies were performed on DNA because of the difficulties of RNA (primarily the difficulties in avoiding degradation caused by RNases), studies on the melting behaviour of RNA as a function of pressure are required to provide more direct information on the physical limitations of RNA under high pressure conditions.

4.6.4 Salts

As observed by Wu and Macgregor [72, 73], the salt concentration present in the DNA solutions changes the magnitude of the effect of pressure on the melting point of DNA sequences. Further experiments should be performed to examine whether this same effect can be observed for the DNA hexamer, particularly since it shows the opposite behaviour to what has already been observed.

Chapter 5 - Physical Stability of Dodecamers

5.1 Introduction

The results of **Chapter 4** have provided some of the first results for the effects of pressure on the melting point of a small DNA sequence. These results, while important, are somewhat limited in their significance. They currently only reflect the effects on a very short sequence, which is largely influenced by the end effects of the terminal bases. The success of the previous studies has prompted for the continuation of these experiments using longer DNA sequences to see if the same, or different, behaviour is observed.

A longer sequence will have two major benefits. First, if the sequence is sufficiently long, the effect of pressure on the melting point of the DNA will be less affected by the instability of the terminal bases. Rather, the melting point will be governed more by the behaviour of the compressibility of the individual nucleotides and the surrounding water molecules. This will allow for a deeper understanding of exactly what forces are affecting the melting point along with being able to see if there are any trends with the length of the DNA.

The second benefit is that, with a longer DNA sequence, there are more possibilities to change the nature of the sequence, primarily, the G-C content. Being able to have different G-C contents in multiple sequences will allow the study of how the two different base pairs, A-T versus G-C, affect the dependency of the physical stability on pressure. This is a very important aspect as changes in the DNA sequence could lead to significant changes in behaviour.

The following is a comprehensive study into the effects of high-pressure on the melting points of five dodecamer DNA sequences where these sequences possess different G-C contents, ranging from 16.7 % to 83.3 %, while maintaining the same terminal bases. Melting experiments were carried out with the use of ^{1}H NMR

spectroscopy and melting points determined by monitoring the chemical shifts of the nucleobase proton signals.

5.2 Materials and Methods

5.2.1 Sample Preparation

Five self-complementary DNA sequences were purchased from Integrated DNA Technologies. These DNA sequences contain varying percentages of G-C content while maintaining the same two terminal bases (5′-C-A .... T-G-3′). These sequences are displayed in Table 5.1.

Sequence	# G-C pairs	G-C %	Final concentration /mM	Reference number
5′-CAT TTA TAA ATG-3′	1	16.7%	0.40	121
5′-CAT TCT AGA ATG-3′	2	33.3%	0.35	122
5′-CAA GTC GAC TTG-3′	3	50%	0.18	123
5′-CAG GTC GAC CTG-3′	4	66.7%	0.46	124
5′-CAC CCG CGG GTG-3′	5	83.3%	0.28	125

Table 5.1: Dodecamer Sequences. Sequences of dodecamers used along with GC content, concentrations used and reference number.

The DNA sequences, whilst cleaned, still contained impurities from the manufacturing process (particularly the protecting group benzamide) which are strongly visible in the ^{1}H NMR spectrum. To remove these impurities, each DNA sequence was first purified by precipitation in cold ethanol.

To maximise the NMR signals, all DNA remaining after purification was used. The final concentration for each DNA sequence is shown in Table 5.1. Only two DNA solutions had concentrations which differed significantly, 123 and 125, and even those are only approximately 50 % and 25 %, respectively, lower than the average concentrations of the other three sequences. While DNA concentration does have an effect on the melting point, it is expected that this level of difference in concentration will have only a small effect on the melting point and no real change on the effects of pressure.

Each DNA sequence was then made up into solution as follows:

A pH 7 0.20 M phosphate buffer solution was first prepared by adding 0.20 M disodium hydrogen phosphate stock solution to 0.20 M monosodium dihydrogen phosphate stock solution (Table 5.2). The disodium hydrogen phosphate was added in the presence of a pH meter to ensure the correct pH of 7.00.

The DNA was dissolved in 100 µL of milli-Q water and combined with, 150 µL of the pH 7 0.10 M phosphate buffer, 30 µL of 5 mM TMSP-d4 stock solution (NMR reference) and 60 µL D_2O (MagniSolv). The solution was made up to 600 µL using milli-Q water, giving final concentrations of 25 mM phosphate buffer and 0.25 mM TMSP-d4 in H_2O/D_2O (90:10 v/v).

Stock Solution	**Mass used /g**	**Volume used /mL**	**Concentration /M**	**Molar Mass /gmol^{-1}**	**Supplier**
$NaH_2PO_4.2H_2O$	1.5601	50.0	0.100	156.01	Riedel-de Haën
$Na_2HPO_4.2H_2O$	1.7799	50.0	0.100	177.99	Riedel-de Haën
TMSP-d4	0.0073	10.00	0.0050	156.26	Merck

Table 5.2: Table of Stock Solutions for Dodecamer Melting Experiments. Each entry shows concentration, mass of compound used, volume made (in milli-Q water) and supplier.

5.2.2 DNA Analysis

Circular dichroism (CD) melting experiments were performed on an Applied Photophysics Chirascan CD spectrometer. A 0.1 mm sample cell was used to reduce the path length through the sample to prevent over saturation without diluting the sample. Melting experiments were performed by recording multiple CD spectra over a temperature range of 278.2 K to 353.2 K in 3 K steps. A temperature equilibration time of 2 minutes was used.

NMR spectra for each DNA sequence were collected with a Bruker Avance 500 MHz NMR spectrometer. Excitation sculpting set at the water offset frequency was used to suppress the dominant water peak in the spectrum.

2D spectra for sequence determination were recorded at 277.2 K using a QXI 5 mm ^{1}H-detection inverse probe with an *z*-field gradient coil. For these initial experiments used for ^{1}H NMR spectrum assignments, each solution was placed in a standard 5 mm glass NMR tube. This was done to increase the sample volume to give a better signal-to-noise ratio.

DNA melting experiments were performed by recording multiple NMR spectra at different temperatures using a TXI 5 mm ^{1}H-detection inverse probe with a z-field gradient coil. For the determination of melting points under high-pressure, each DNA solution was transferred to the high-pressure NMR cell. NMR spectra were recorded with 3 K temperature steps over a temperature range of at least 39 K. A time of 5 minutes between each step was allowed for temperature equilibration. All 1D spectra were recorded with 512 scans, relaxation time of 2 s, spectral width 20.7 ppm and 65536 points. All 2D spectra were recorded with 512 scans, relaxation time of 2 s, a recycle delay of 2 s and 4096 x 512 points. The COSY spectral width was 18.9 x 19.0 ppm while the NOESY spectral with was 18.9 x 10.2 ppm and mixing time of 0.5 s.

5.2.3 Spectral Assignment of Sequence 123

The spectral assignments of each of the five DNA sequences were carried out using the same method as that used for the assignment of 5′-CAT ATG-3′ in **Chapter 4**. The following is a brief reminder of this method mentioning any difference in procedure. The same signal and sequence position naming nomenclature used during **Chapter 4** is also used for this Chapter.

To obtain appropriate NMR spectra for spectral assignment for each DNA sequence, the following method was used.

- A brief melting experiment was first carried out to obtain two key temperatures. First, obtain an initial indication of the DNA's melting point. This is used to determine the central temperature for the melting range for each sequence. Secondly and more importantly for the sequencing, to determine a temperature, T_s, at which there is the greatest degree of spectral separation (lowest number of overlapped signals) in the region of interest, 8.5 ppm to 6.9 ppm. This is to ensure that there is sufficient separation of NMR signals to reduce the number of overlapping peaks in the 2D spectra.

- Both NOESY and COSY 2D NMR spectra of the DNA strand were recorded at T_s. These spectra were then used in the spectral assignment of each DNA sequence.

Assignments were made using the method outlined in **Chapter 4**. To display the increase in complexity that simply doubling the length of the sequence can provide, an example of the assignment of a dodecamer sequence is provided.

The assignment of sequence 123 (5′-CAA GTC GAC TTG-3′) is as follows:

With the use of Figure 4.2, Figure 4.3 and a NOESY spectrum, the type of base for each signal can be identified (Figure 5.1). The C-6H and A-2H peaks can be identified first. Cytosine signals show unique doublets while A-2H peaks exhibit no NOE's with 2′H/2″H protons. Once these peaks are eliminated the remaining

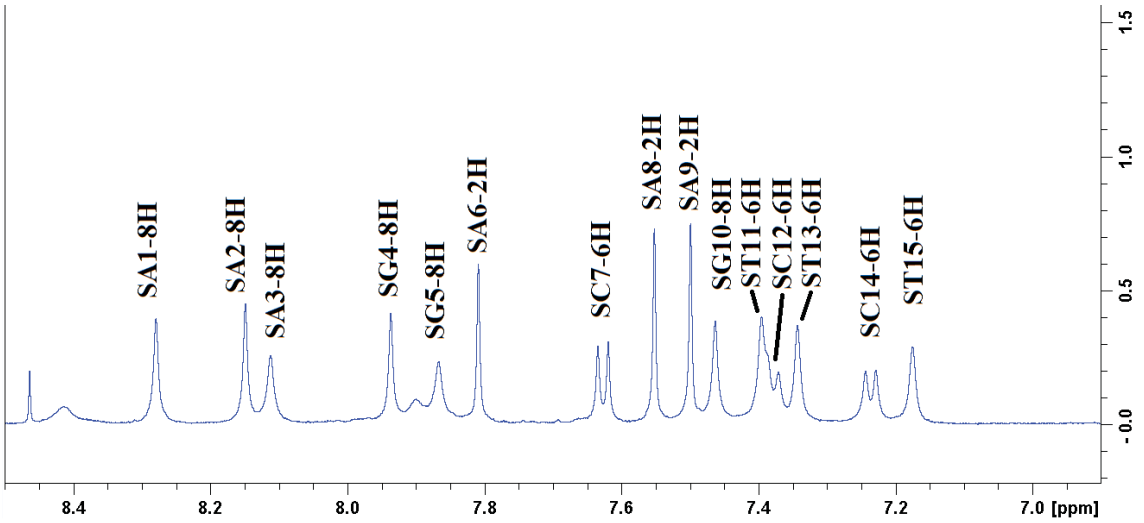


Figure 5.1: ^{1}H NMR Spectrum of Sequence 123 (5′-CAA GTC GAC TTG-3′). Protons have been assigned as described in the text.

signals can be identified by comparing the position of each peak with the spectral ranges shown in Figure 4.2.

The terminal bases can now be identified. The 5′-terminal cytosine can be identified by comparing the observed signals in the region of 8.5-6.9/6.5-5.5 ppm. It can be seen that (with the exception of the C-6H→C-5H signal) S7C-6H has no vertical links from the intranucleotide C-6H→C-1′H signal (Figure 5.2). This indicates S7C-6H as the terminal cytosine. Similarly, 3′-terminal guanine can be identified as the guanine since there is no horizontal link from the intranucleotide NOE cross peak. It can be seen in Figure 5.2 that this is S4G-8H.

Three links can be clearly identified from the CH_3 groups on the thymine residues between the cross peaks: S10G-8H→CH_3 – CH_3→S15T-6H, S14C-6H→CH_3 – CH_3→S11T-6H and S11T-6H→CH_3 – CH_3→S13T-6H (Figure 5.3). These account for connectivities between G4-8H→T5-6H, C9-6H→T10-6H and T10-6H→T11-6H. Since there is only one of each of the GT and CTT sequences, it is clear that these signals belong to these sections of the sequence.

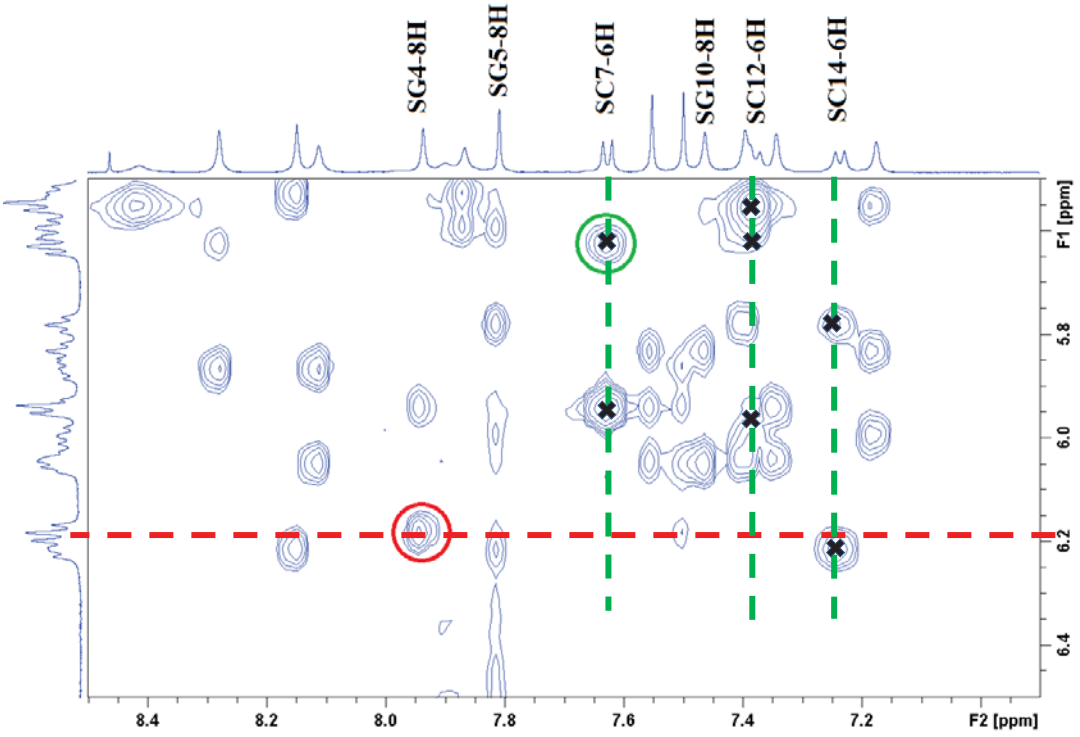


Figure 5.2: Identification of Terminal Bases. The 5′-terminal cytosine base (green circle) and the 3′-terminal guanine deoxyribose (red circle). For the three cytosine signals (green dashed lines) it can be seen that there are two which have three vertical links while S7C-6H only has two (the third signal for S14C-6H is located at about 5.2 ppm and is not visible in this image). This is a C-6H→C-5H signal and is not relevant to the assignment of this sequence. After excluding the C-6H→C-5H signal, S7C-6H has only one link and is therefore the 5′-terminal base. Similarly for guanine, it can be seen that there is no horizontal link from the intranucleotide NOE on S4G-8H (red dashed line), identifying it as the 3′-terminal base.

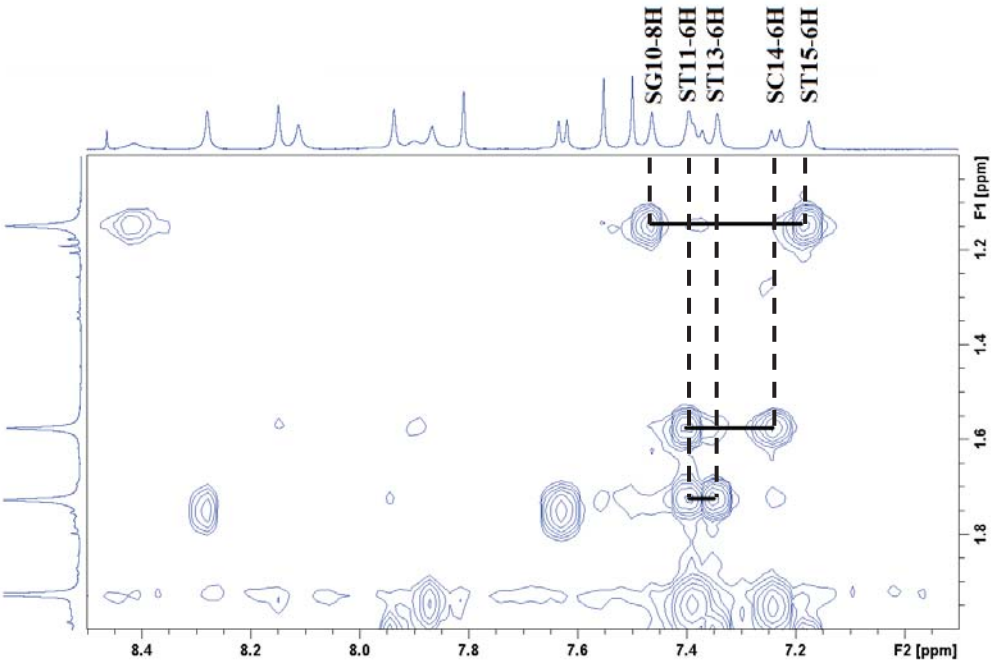


Figure 5.3: S10G-8H→CH_3 – CH_3→S15T-6H, S14C-6H→CH_3 – CH_3→S11T-6H and S11T-6H→CH_3 – CH_3→S13T-6H Links. Links are indicated by solid horizontal lines.

Using the information above, an assignment link pathway using d_i(6H/8H→1′H) and d_s(1′H→6H/8H) connectivities can be determined (Figure 5.4) and the DNA sequence can be assigned as follows (Table 5.3).

Base	C1	A2	A3	G4	T5	C6
Signal	SC7-6H	SA1-8H	SA3-8H	SG10-8H	ST15-6H	SC12-6H
Chemical shift at 305.2 K	7.627	8.280	8.112	7.465	7.175	7.381
Base	G7	A8	C9	T10	T11	G12
Signal	SG5-8H	SA2-8H	SC14-6H	ST11-6H	ST13-6H	SG4-8H
Chemical shift at 305.2 K	7.869	8.150	7.237	7.396	7.343	7.937

Table 5.3: Assignment of Sequence 123.

This result is confirmed with d_i(6H/8H→2′H)/ d_i(6H/8H→2″H) and d_s(2′H→6H/8H)/d_s(2″H→6H/8H) sequential link pathways with both NOESY and combined NOESY-COSY link diagrams, as shown in Figure 5.5 and Figure 5.6.

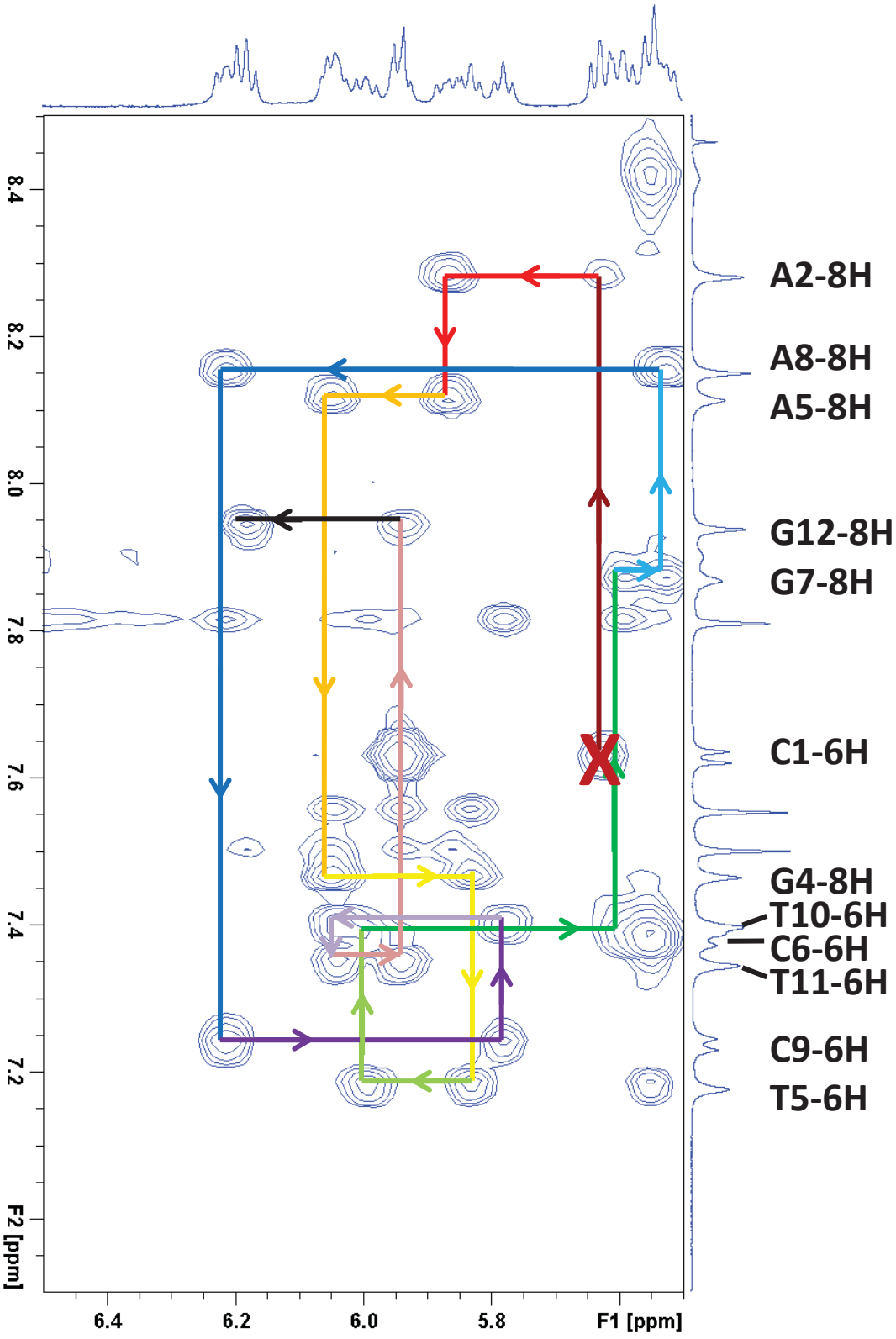


Figure 5.4: NOESY Assignment Link Pathway for Sequence 123 (5′-CAA GTC GAC TTG-3′) with d_i(6H/8H→1′H) and d_s(1′H→6H/8H) Connectivities. The link pathway can be followed from the dark red X along the lines in "rainbow" order through to the final black line.

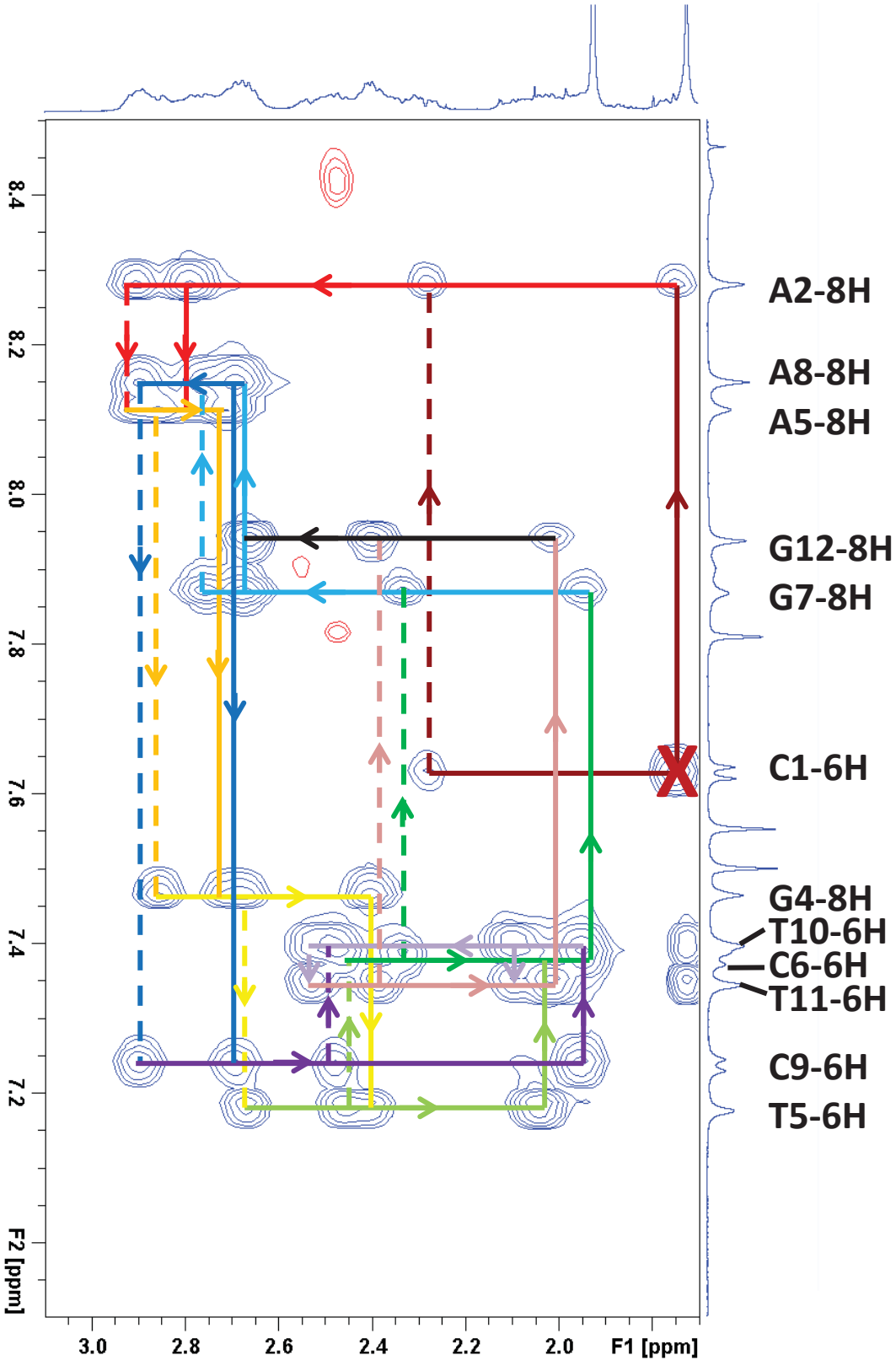


Figure 5.5: NOESY Assignment Link Pathway for Sequence 123 (5′-CAA GTC GAC TTG-3′) with d_i(6H/8h→2′H)/ d_i(6H/8H→2″H) and d_s(2H′→6H/8H)/d_s(2″H→6H/8H) Connectivities. Solid horizontal lines indicate 2′ i-s links and horizontal dashed lines represent 2″ i-s links (as seen side on). The link pathway can be followed from the dark red X along the lines in "rainbow" order through to the final black line.

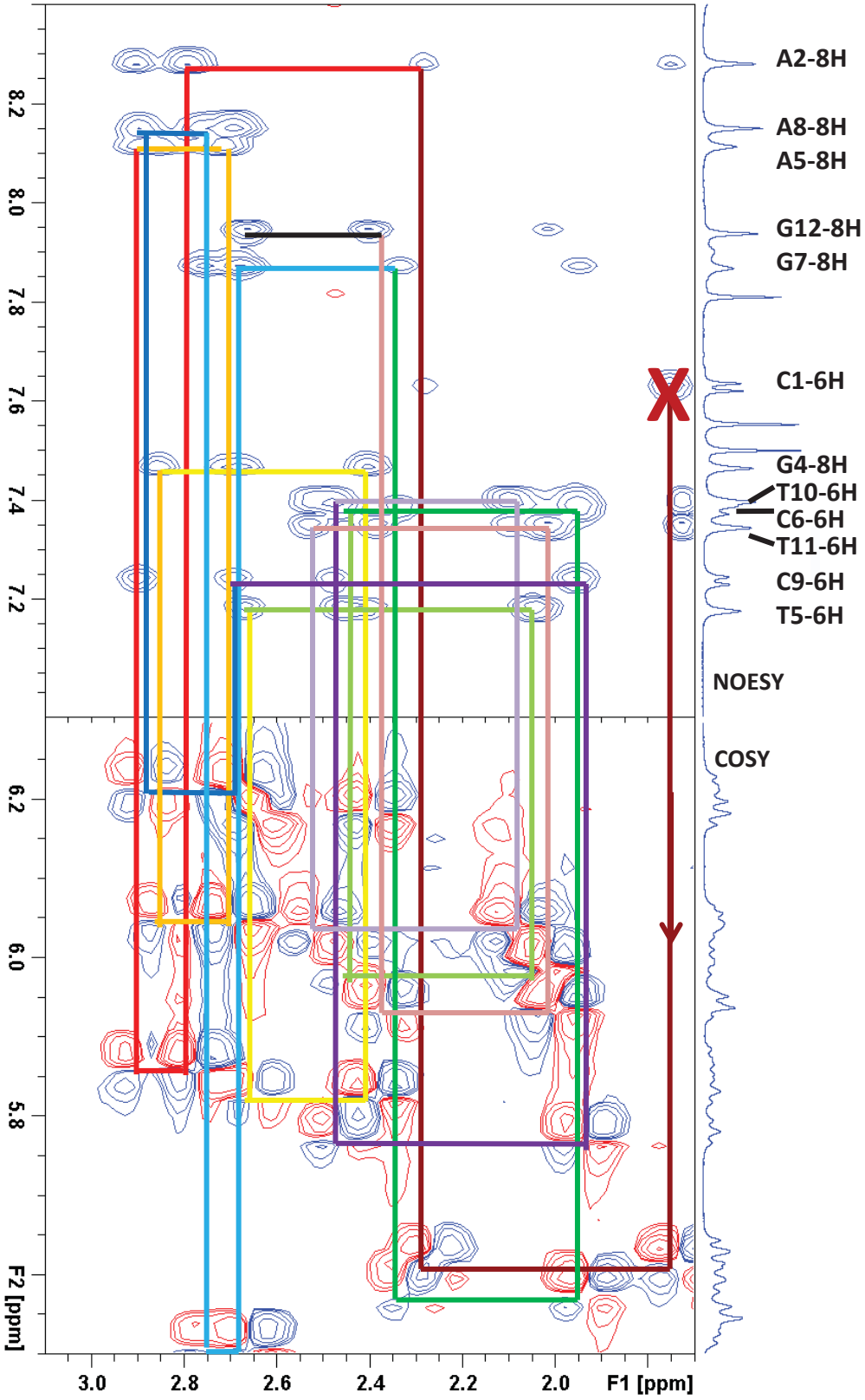


Figure 5.6: Combined NOESY-COSY Assignment Link Pathway for Sequence 123 (5′-CAA GTC GAC TTG-3′) with d_i(6H/8H→2H′) and d_s(2″H→6H/8H) Connectivities. The signals forming the COSY spectrum are each comprised of two symmetrical cross-peaks (one positive, blue, and one negative, red), each along the one of the diagonals of the spectra. These cross peaks are a result of magnetisation transfer between two coupled protons. Where these diagonals cross is the centre of the overall signal. Note: some of these cross-peaks overlap or are not fully visible at this resolution. The link pathway can be followed from the dark red X along the lines in "rainbow" order through to the final black line.

5.2.4 Melting Experiments

After successfully assigning the ^{1}H NMR spectra, high-pressure melting experiments were performed for each DNA sequence. These experiments involved heating of the sample in 3 K stepped temperature increments while recording a ^{1}H NMR spectrum at each step. This process was repeated at pressures of 0.1, 50, 100, 150 and 200 MPa using the high-pressure cell. These NMR spectra were then used to determine the melting point for the DNA sequences under different pressure conditions. Melting experiments done at 0.1, 50, 100 and 200 MPa had only single runs while experiments done at 150 MPa were run in triplicate.

5.2.5 Circular Dichroism Experiments

A circular dichroism melting experiment was performed in order to obtain an accurate melting point to confirm the results obtained using the ^{1}H NMR spectroscopy. A 50 µL drop of the DNA solution was placed in a 0.1 mm quartz cell. The CD spectrum of the solution was recorded at 3 K stepped temperature increments at 0.1 MPa. These spectra were then used to determine the melting point of the DNA to compare to the results from the NMR spectroscopy.

5.3 Results

5.3.1 Data Gathering and Processing

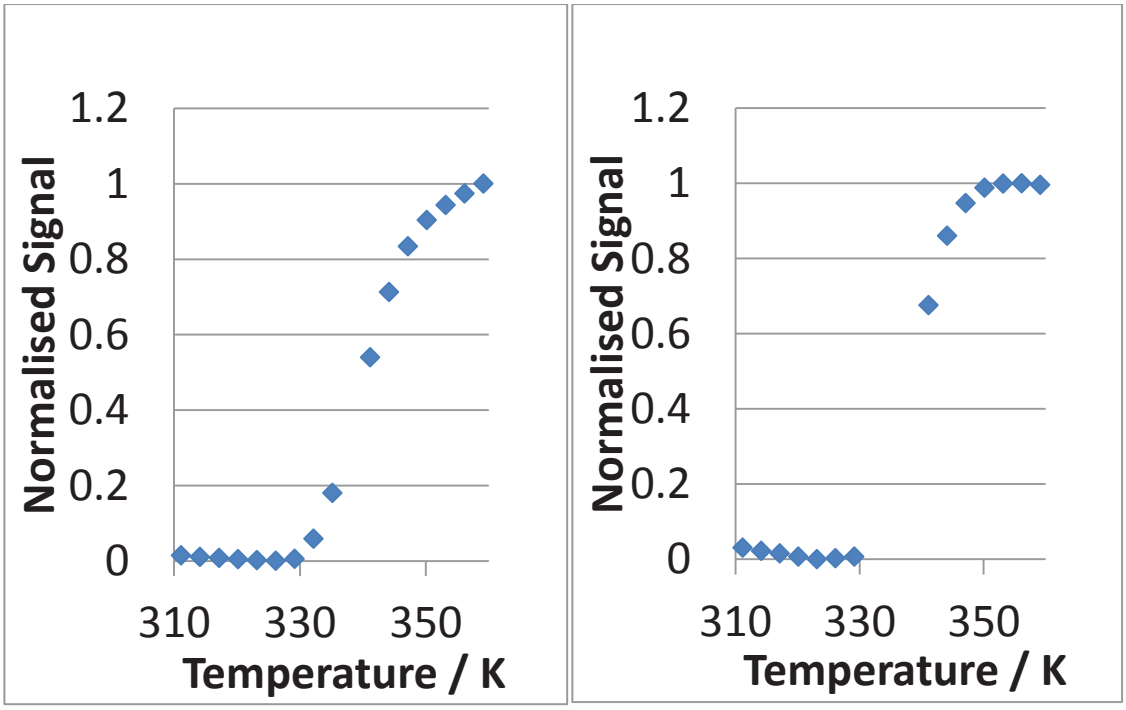
The following was repeated for each DNA sequence.

For each spectrum obtained at 0.1 MPa, the chemical shifts for the signals corresponding to protons C-6H, A-8H, T-6H, and G-8H were recorded. This

involved a process of recording the chemical shift of the proton signals through the melting of the DNA.

Due to the signal broadening and movement, each signal had to be carefully monitored through the melting process so as to obtain the correct curve. For many of the signals this was not possible. In these instances, the corresponding signal after melting was determined and the progression of the signal was followed back through the melting process until it was lost again. This process made it possible to determine the full curve for most of these signals; however, approximately 10% of all the melting curves lacked sufficient data points for correct fitting (Figure 5.7). The melting points obtained for these incomplete melting curves can be considered estimates at best and, therefore, have been treated with caution. This process was repeated for all melting data acquired for each pressure condition.

Melting curves for each residue for each sequence are given in Figure 5.8 – Figure 5.12.



(a) (b)

Figure 5.7: Examples of Melting Curves Where Data Gathering Proved Difficult. a) An incomplete curve where the signal has been followed from either end of the melting process to establish an almost full curve. b) A curve where the missing signals in the melting region have left insufficient data for accurate fitting of the melting curve.

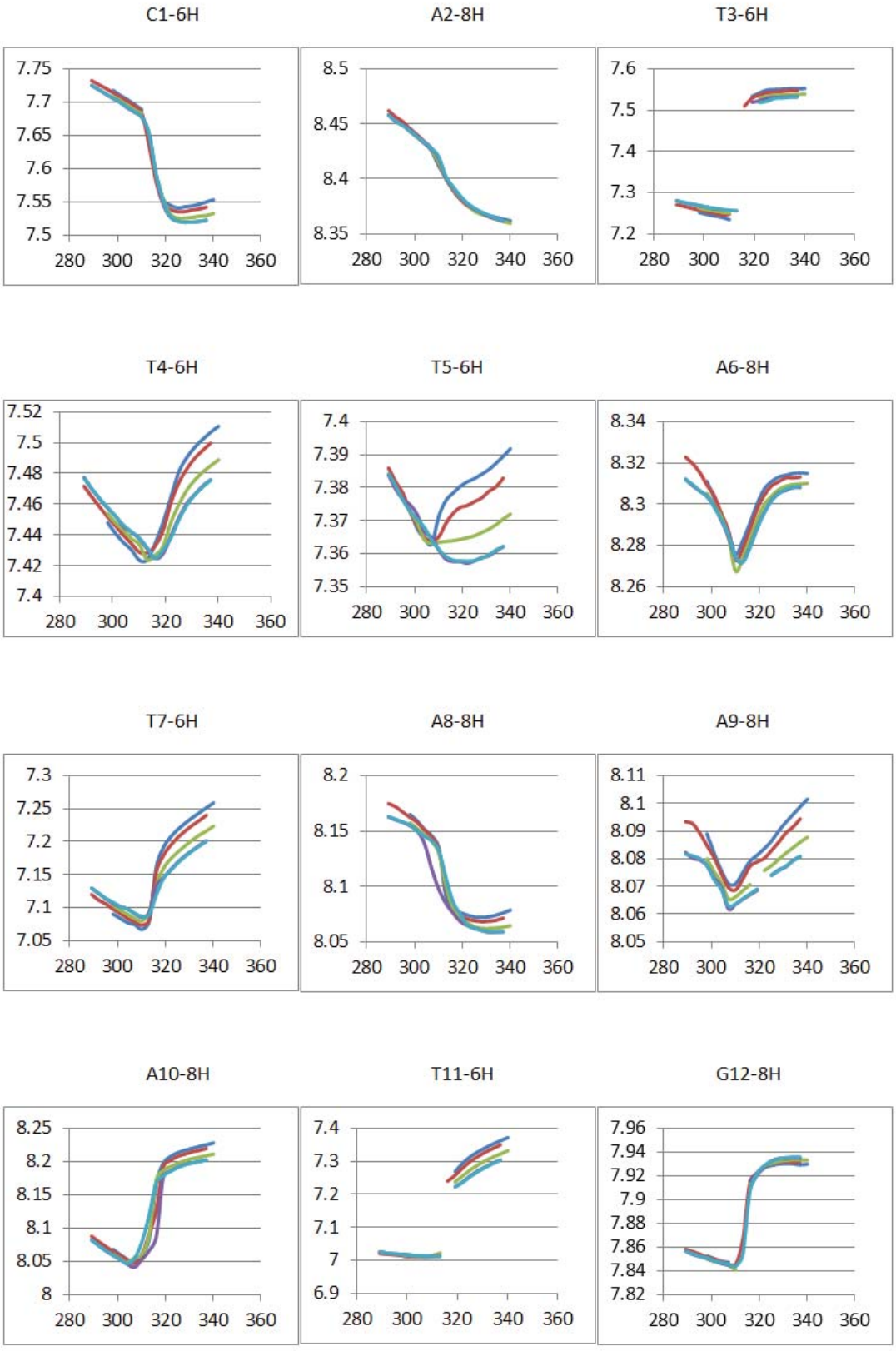


Figure 5.8: Residue Melting Curves for Sequence 121. 0.1 MPa – blue, 50 MPa – red, 100 MPa – green, 150 MPa – light blue, purple. Melting curves are presented as the chemical shift (/ppm) as a function of temperature (/K).

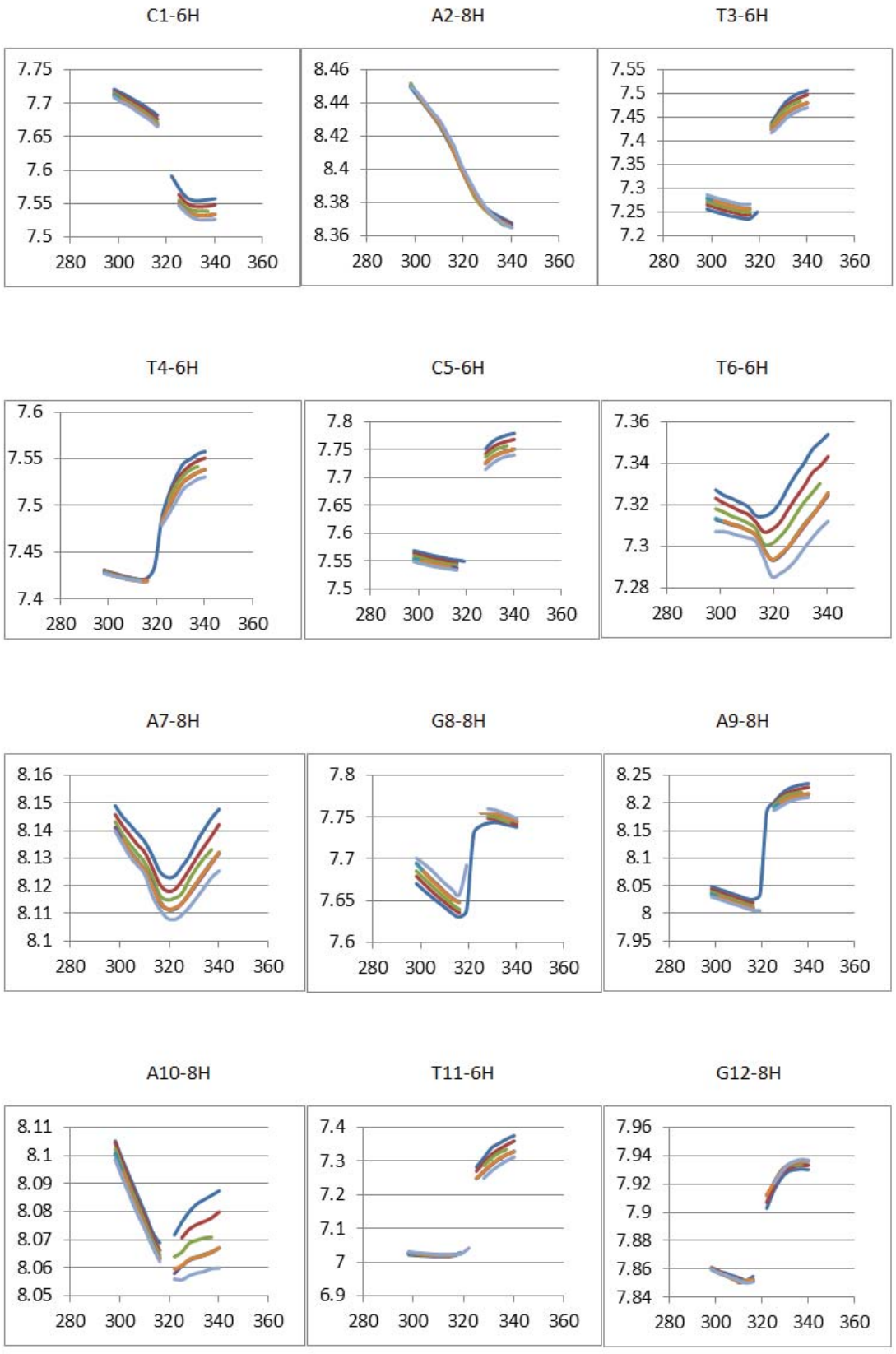


Figure 5.9: Residue Melting Curves for Sequence 122. 0.1 MPa – blue, 50 MPa – red, 100 MPa – green, 150 MPa – orange, light blue, purple, 200 MPa – light purple. Melting curves are presented as the chemical shift (/ppm) as a function of temperature (/K).

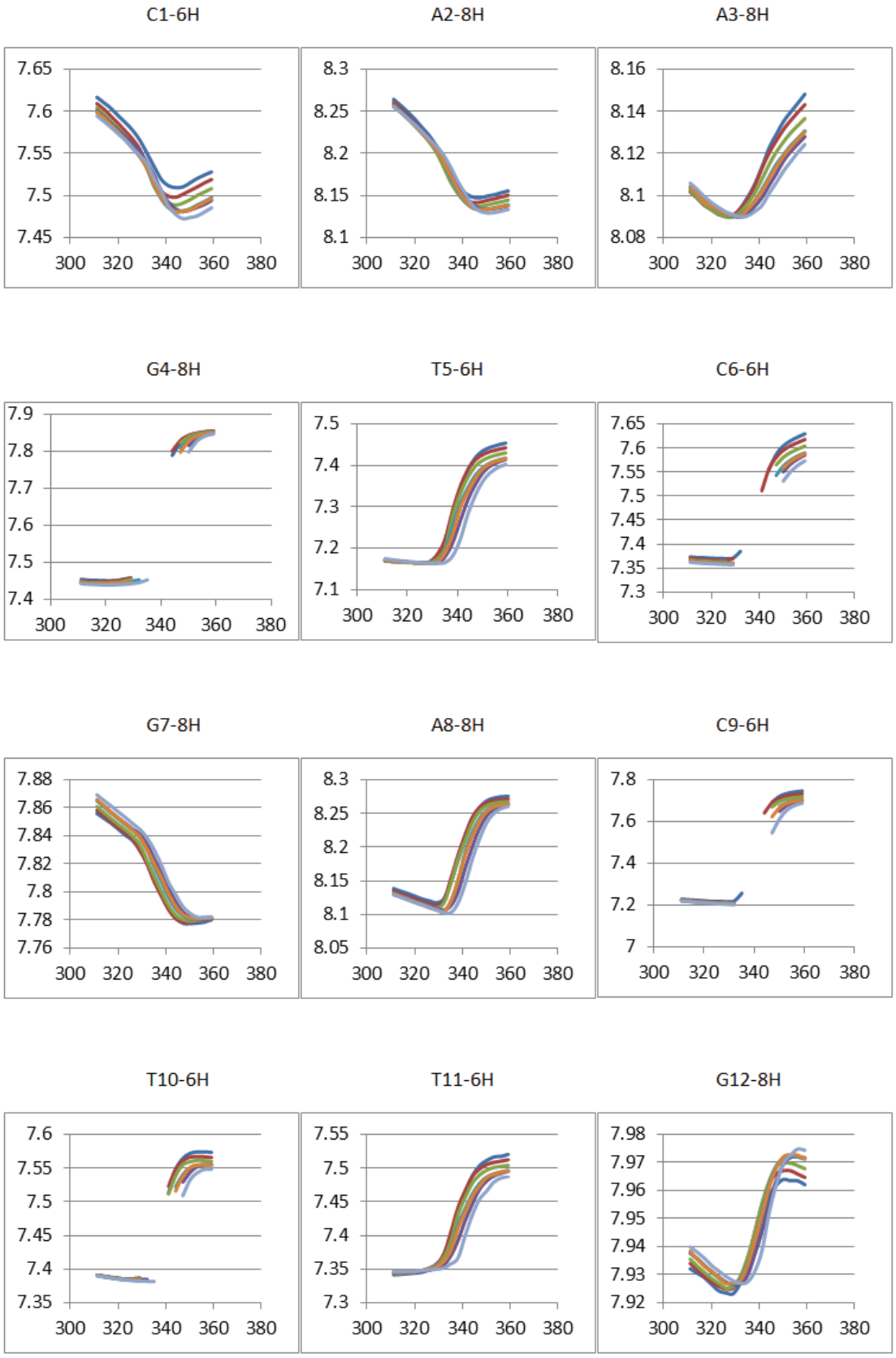


Figure 5.10: Residue Melting Curves for Sequence 123. 0.1 MPa – blue, 50 MPa – red, 100 MPa – green, 150 MPa – orange, light blue, purple, 200 MPa – light purple. Melting curves are presented as the chemical shift (/ppm) as a function of temperature (/K).

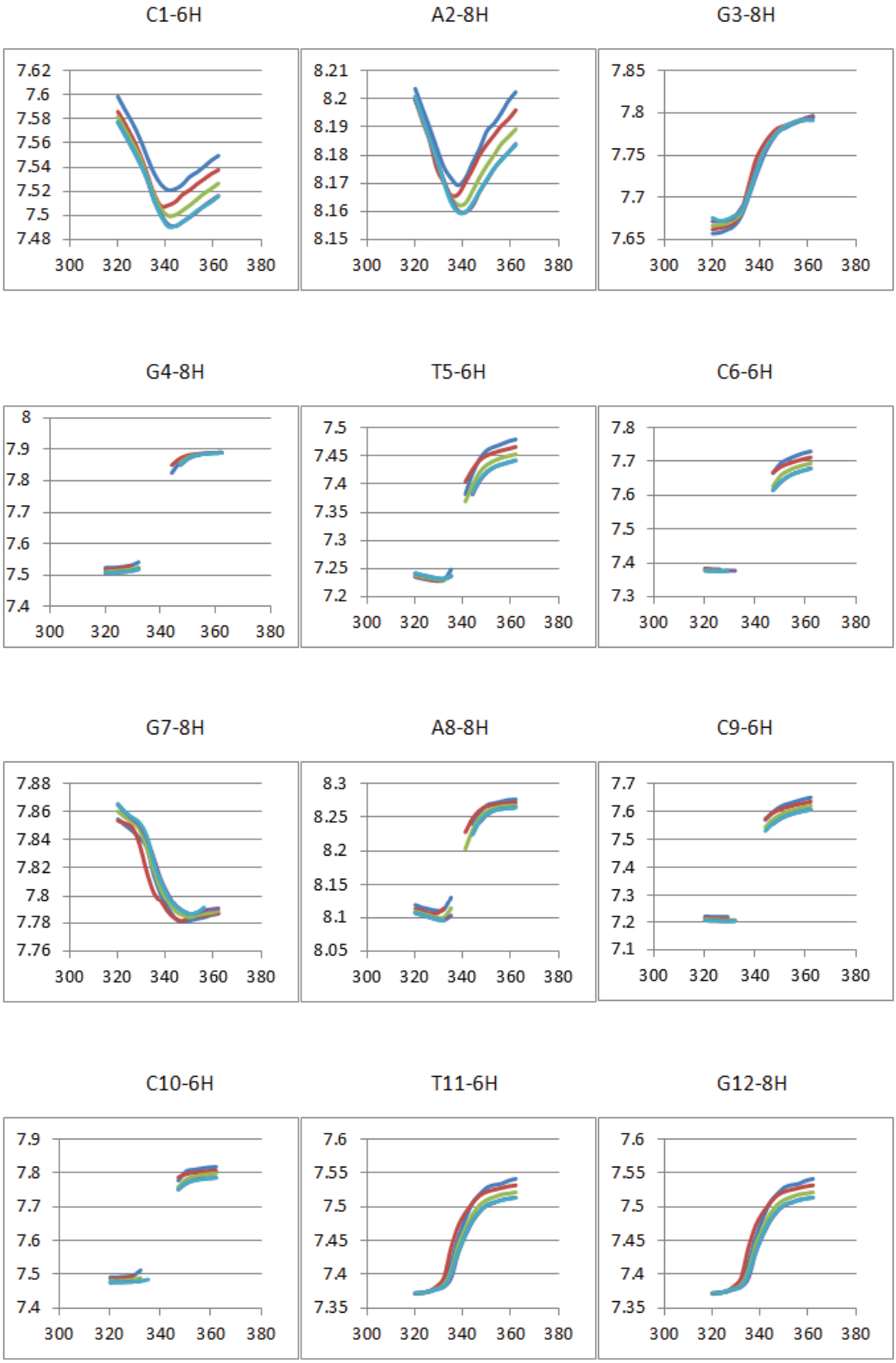


Figure 5.11: Residue Melting Curves for Sequence 124. 0.1 MPa – blue, 50 MPa – red, 100 MPa – green, 150 MPa – light blue, purple. Melting curves are presented as the chemical shift (/ppm) as a function of temperature (/K).

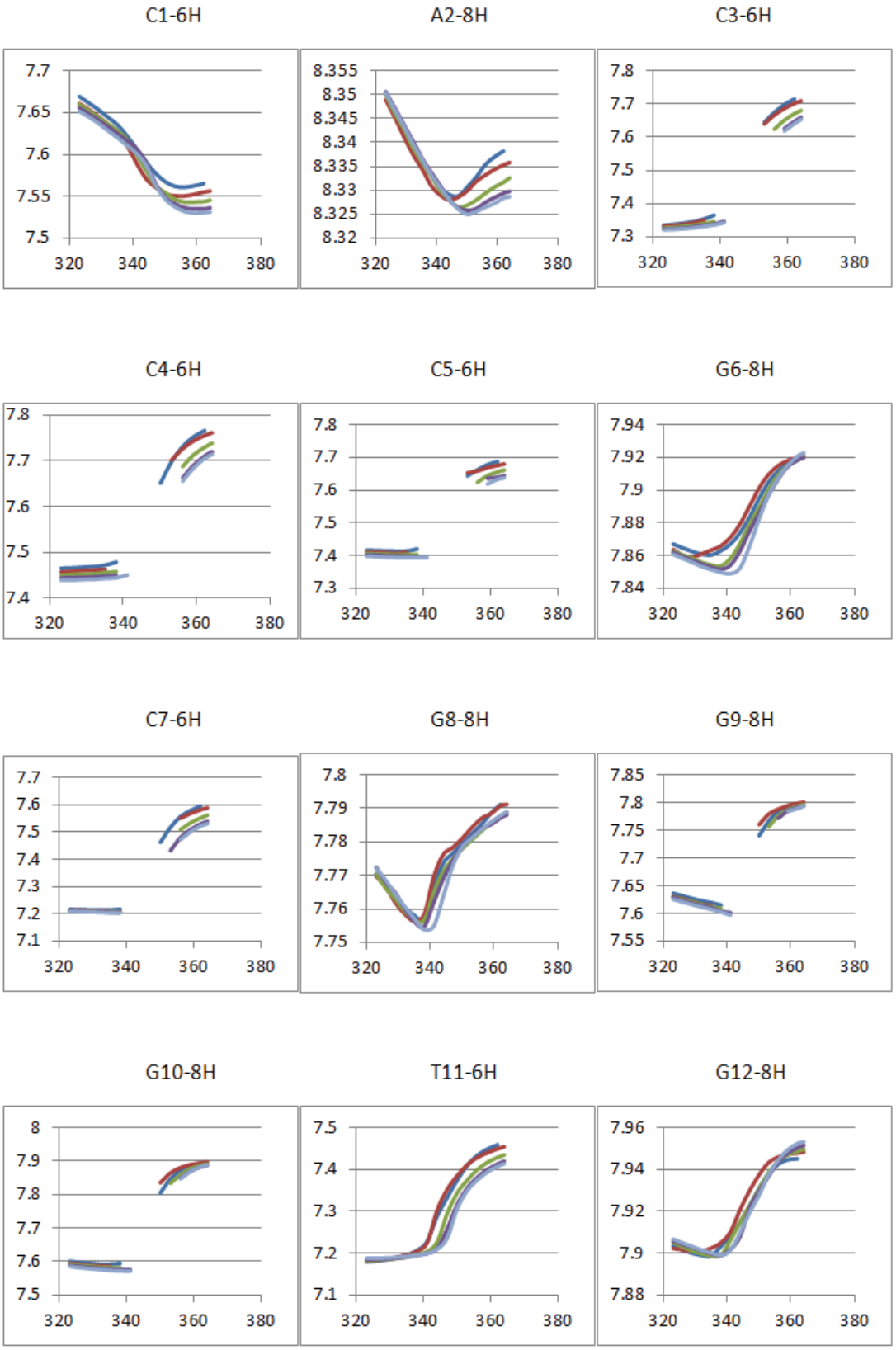


Figure 5.12: Residue Melting Curves for Sequence 125. 0.1 MPa – blue, 50 MPa – red, 100 MPa – green, 150 MPa – light blue, purple, 200 MPa – light purple. Melting curves are presented as the chemical shift (/ppm) as a function of temperature (/K).

5.3.2 Data Fitting

Data fitting followed the same procedure described in **Chapter 4**. The melting curves obtained were fitted using Gnuplot and equation (23) to determine the melting points for each pressure condition and for each DNA sequence. This process proved to be more complex than in **Chapter 4** due to the variation of melting curves. Plots of melting point versus pressure for each residue are presented below for each sequence (Figure 5.13 – Figure 5.17). Fitting parameters are available in the digital Appendices.

For some curves it was not possible to successfully determine the melting point from the melting curve by fitting of equation (23). This was a result of their more complex curves. In these instances, the melting curve was fitted empirically to a polynomial function. The point of inflection in the middle of the curve was then determined by taking the double derivative of this function, giving the melting point. For four residues the shape of the melting curves prevented successful fitting of both equation (23) and the polynomial function. However, for two of these residues (sequence 121 T4-6H and sequence 124 G7-8H), the melting points could be visually estimated as the midpoint of the sigmoidal-shaped section. For the remaining two residues (sequence 124 A2-8H and sequence 125 A2-8H) there was no discernible sigmoidal section and therefore the best fit of equation (23) was used as an estimated melting point.

As mentioned earlier, incomplete melting curves have resulted in estimated melting points. These estimates are highly influenced by the number and location of the missing points on the melting curve leading to an experimental uncertainty of several degrees. This inaccuracy is clearly seen in some of the melting point versus pressure plots as outlying curves. The curves corresponding to these estimated melting points can be excluded from further calculations to give a more accurate representation of the melting point versus pressure for each sequence.

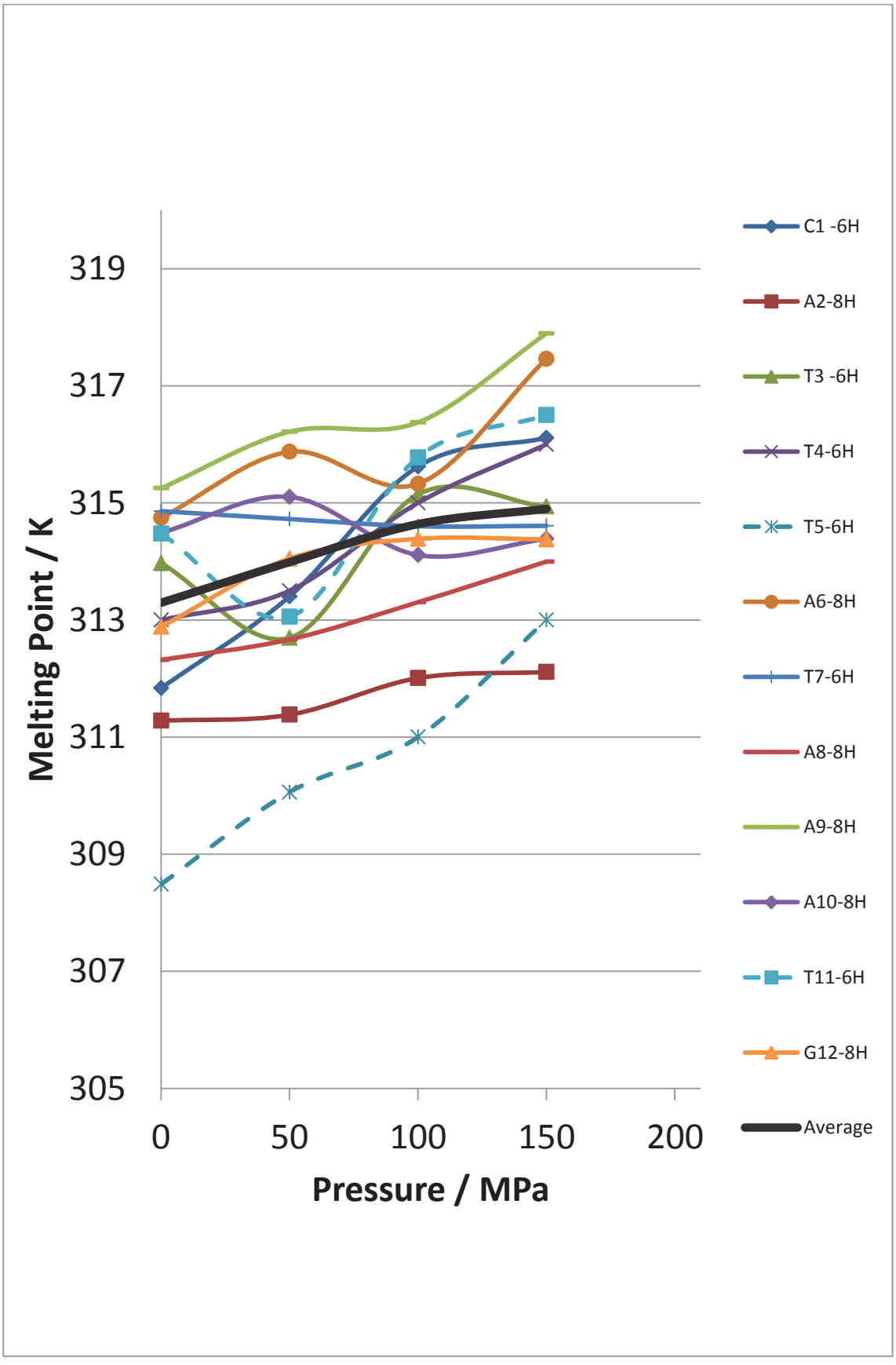


Figure 5.13: Melting Point versus Pressure for Sequence 121 (5′-CAT TTA TAA ATG-3′). Plots for each residue are displayed along with the average (thick black line). Plots where melting points could only be approximately estimated are shown as dashed lines.

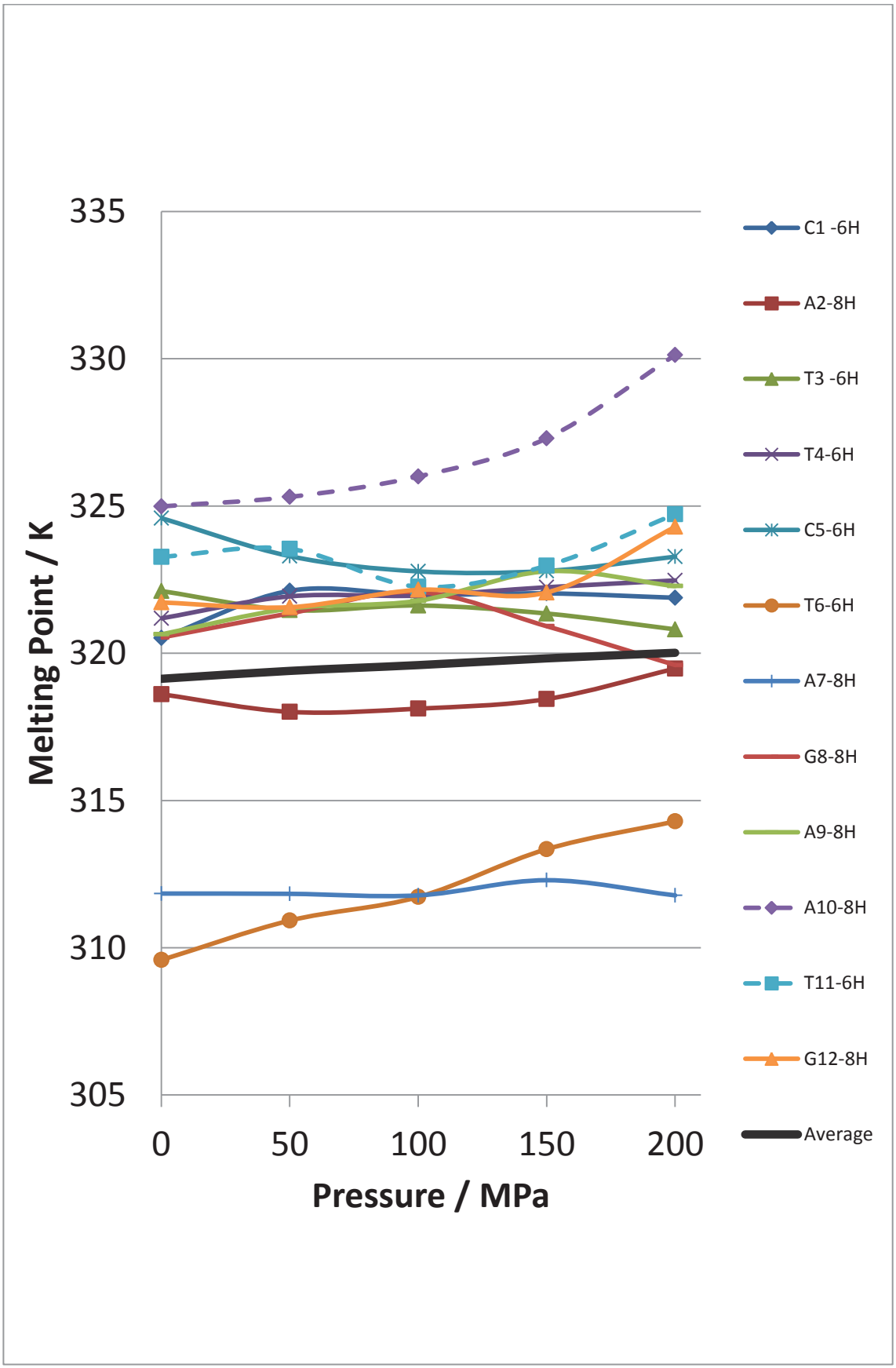


Figure 5.14: Melting Point versus Pressure for Sequence 122 (5′-CAT TCT AGA ATG-3′). Plots for each residue are displayed along with the average (thick black line). Plots where melting points could only be approximately estimated are shown as dashed lines.

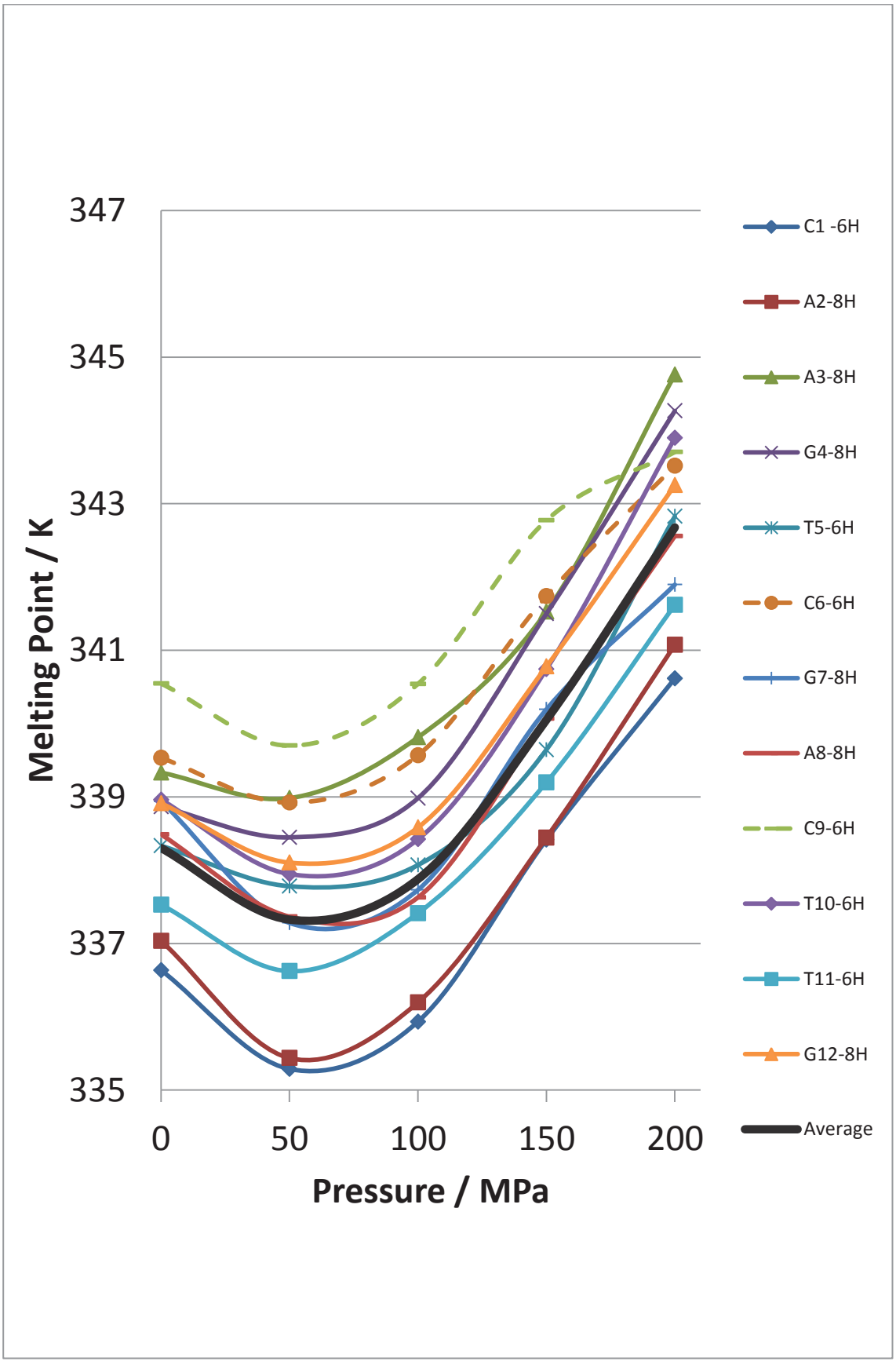


Figure 5.15: Melting Point versus Pressure for Sequence 123 (5′-CAA GTC GAC TTG-3′). Plots for each residue are displayed along with the average (thick black line). Plots where melting points could only be approximately estimated are shown as dashed lines.

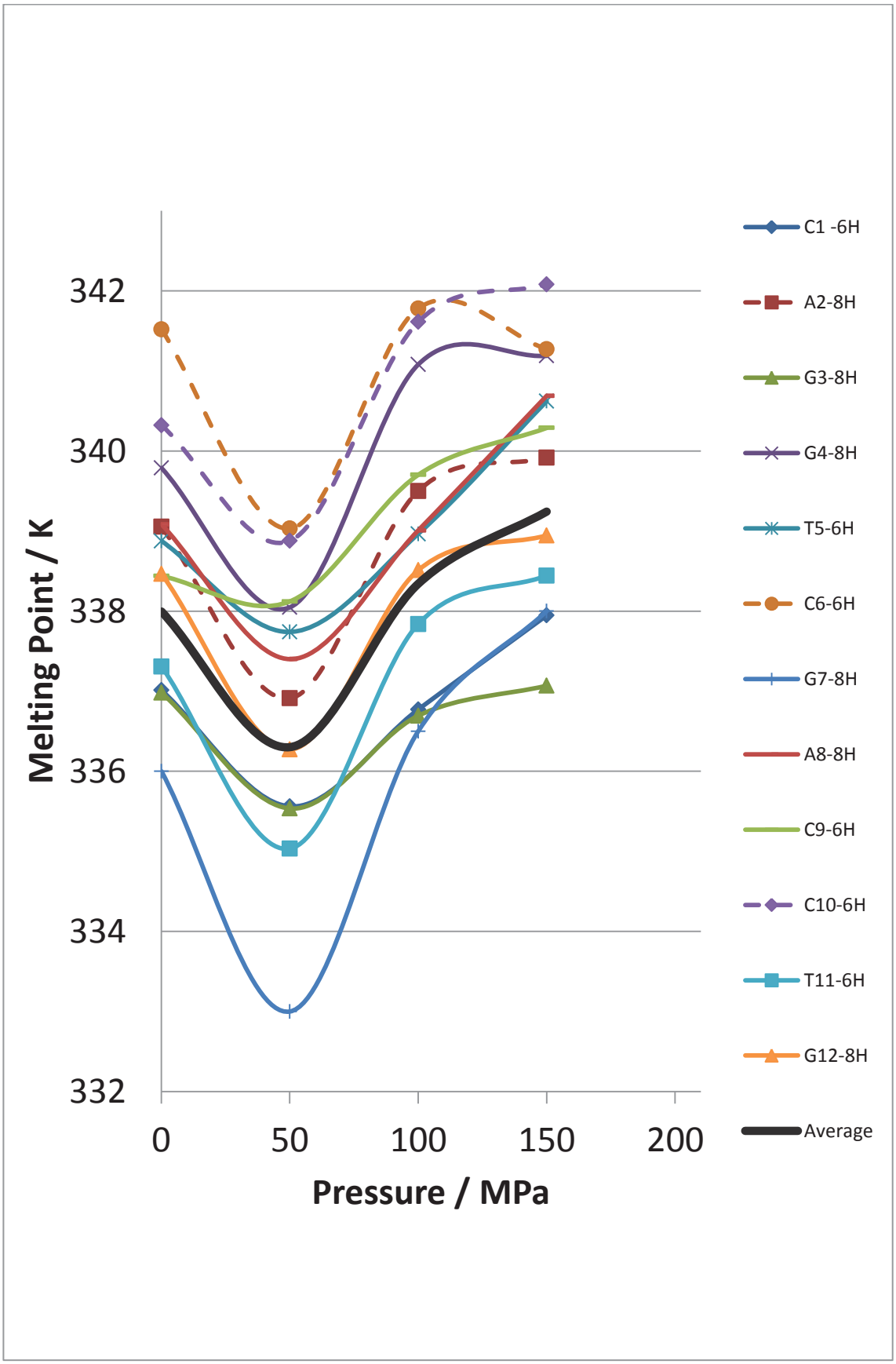


Figure 5.16: Melting Point versus Pressure for Sequence 124 (5′-CAG GTC GAC CTG-3′). Plots for each residue are displayed along with the average (thick black line). Plots where melting points could only be approximately estimated are shown as dashed lines.

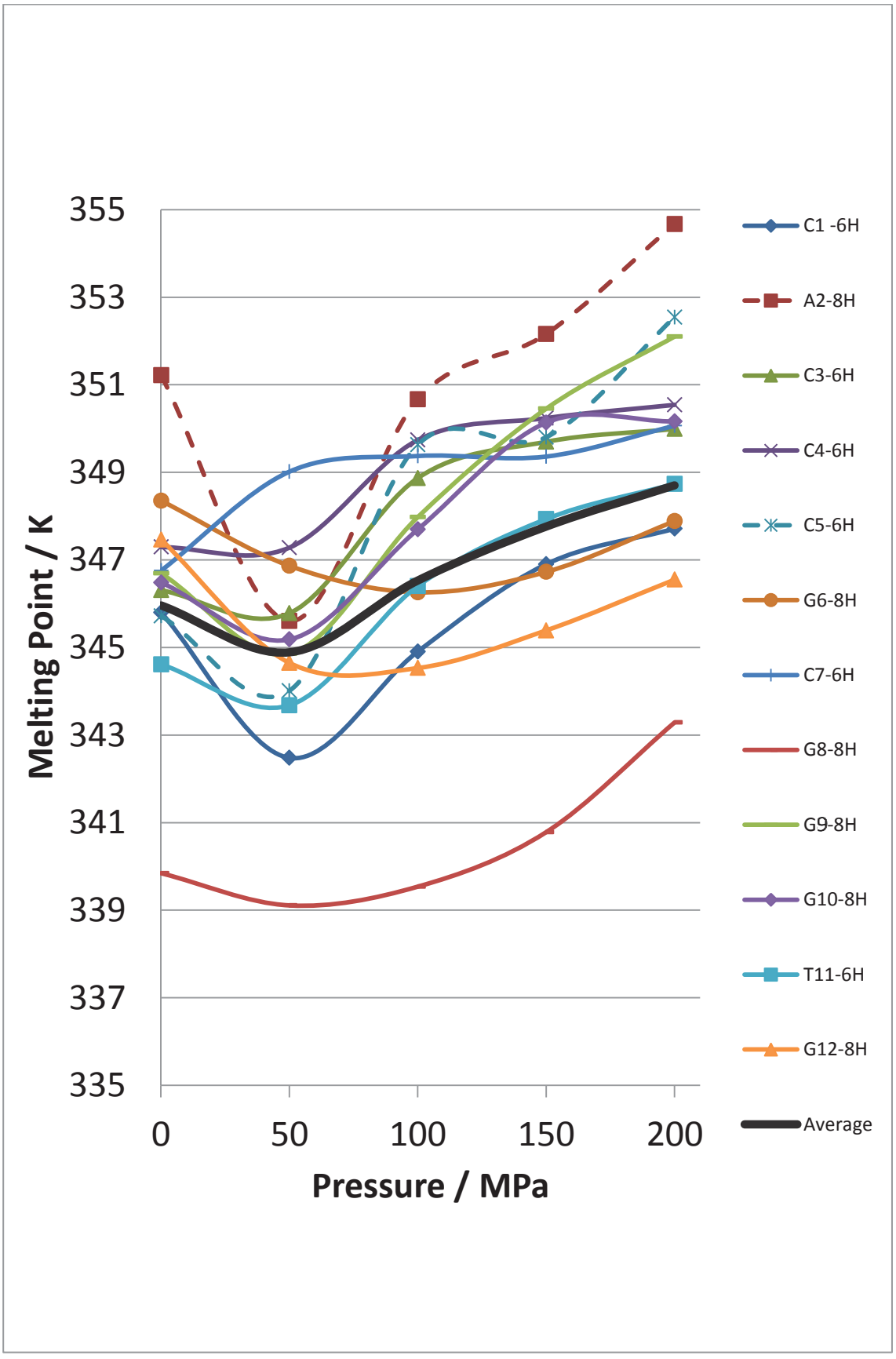


Figure 5.17: Melting Point versus Pressure for Sequence 125 (5′-CAC CCG CGG GTG-3′). Plots for each residue are displayed along with the average (thick black line). Plots where melting points could only be approximately estimated are shown as dashed lines.

5.3.3 Circular Dichroism Results

The melting points of each DNA sequence were also determined by CD spectroscopy. Figure 5.18 shows example CD spectra, obtained for the melting of the DNA sequence 122. The signal intensity's of the peaks in the ranges of: 205 nm – 220 nm, 235 nm – 250 nm and 270 nm – 280 nm, were recorded, smoothed over 3 nm, normalised and plotted as a function of temperature (Figure 5.19) for each sequence. The melting point of each DNA sequence is then determined as the average temperature at intensity 0.5. The results of the melting of each dodecamer DNA sequence are presented in Table 5.4 alongside the average melting points obtained via NMR spectroscopy. The non-perfect isobestic points can be attributed to the formation of small bubbles in the solution during heating.

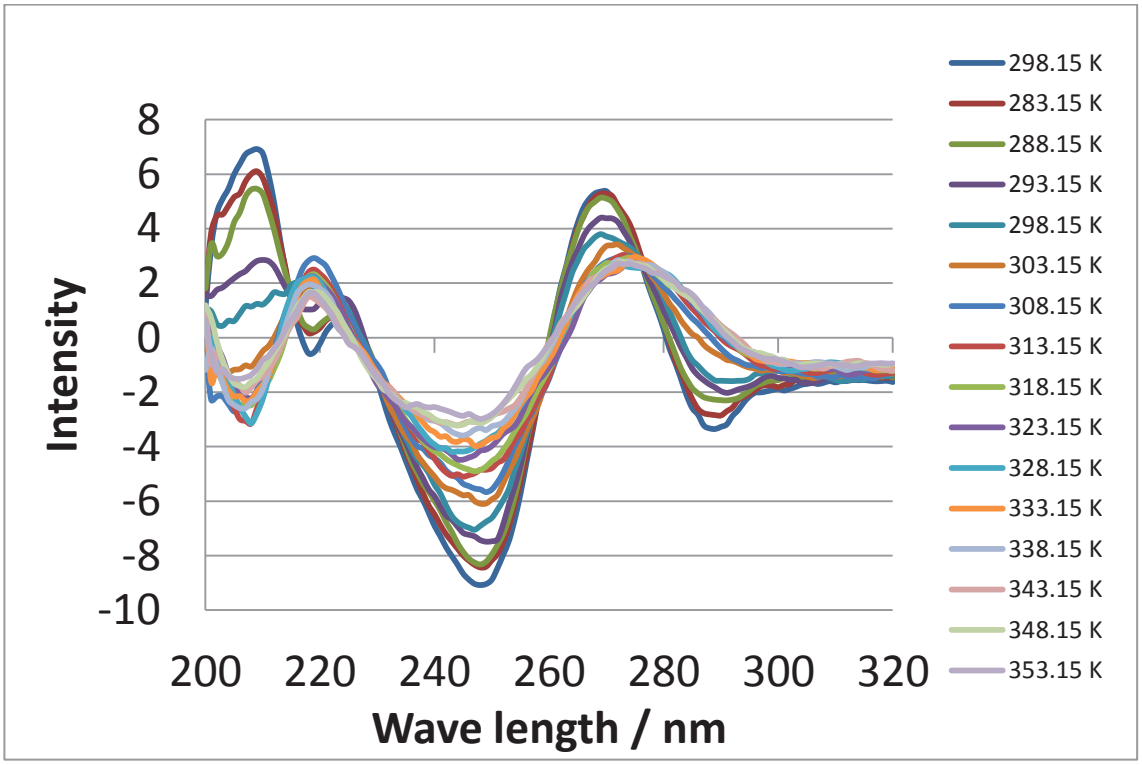
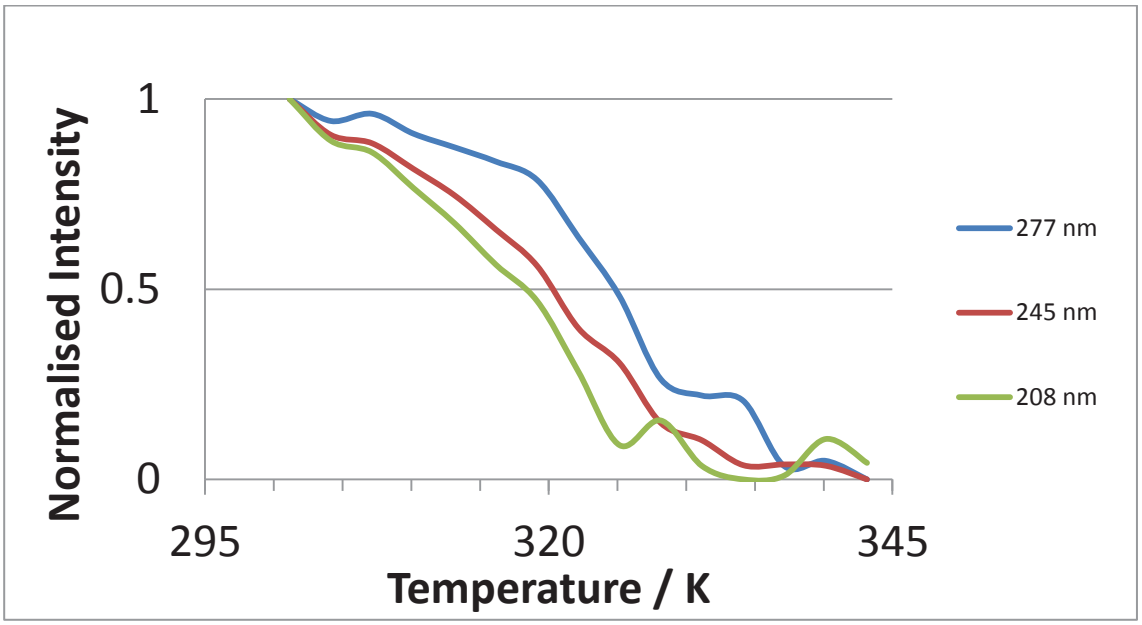
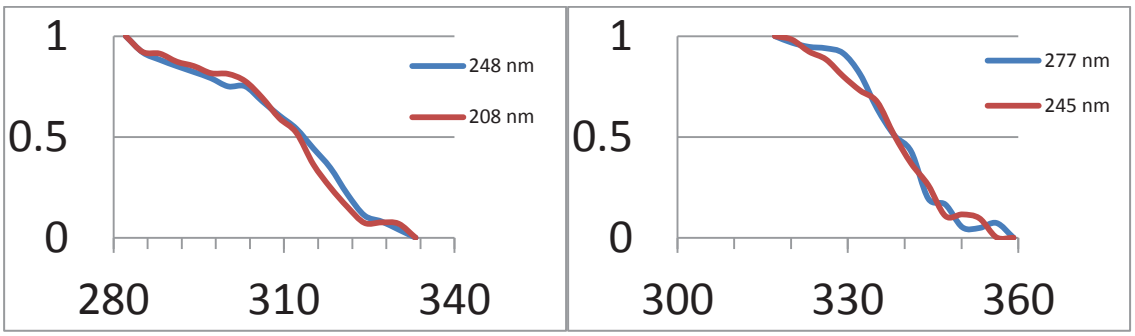


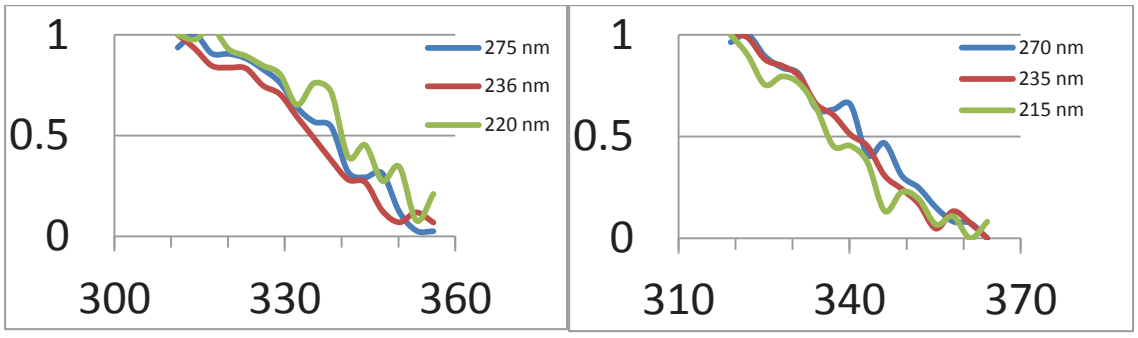
Figure 5.18: CD Spectra of Sequence 122 During DNA Melting.



a)



b) c)



d) e)

Figure 5.19: CD Melting Curves. a) Melting curves obtained via CD spectroscopy at wavelengths of 208, 245 and 277 nm for sequence 122. The melting point can be determined by the temperature corresponding to the average normalised intensity of 0.5. Also presented are sequences b) 121, c) 123, d) 124 and e) 125. CD melting spectra and larger melting curve images are available in Appendix 3.

Sequence	CD T_m / K	Average NMR T_m / K
121	313.2 ± 2	313.5 ± 2
122	321.2 ± 2	320.4 ± 2
123	338.2 ± 2	338.3 ± 2
124	338.2 ± 2	338.6 ± 2
125	343.2 ± 2	346.4 ± 2

Table 5.4: CD Melting Results of Dodecamers. The CD melting temperature results are presented here alongside the results obtained from NMR spectroscopy at 0.1 MPa.

5.4 Discussion

The results of **Chapter 4** gave an interesting insight in to the behaviour of a small DNA hexamer under high-pressure conditions. However, this study was limited to a single short DNA sequence and therefore the results obtained reflect a narrow view of the DNA world. To better understand the effects of the physical environment on the physical stability of a small DNA sequences, the methodology of **Chapter 4** was extended to study the effects of high pressures on the physical stability of DNA dodecamers of varying G-C content to determine whether there are any behavioural trends.

5.4.1 Sequence 121 (16.7 % G-C Content)

Figure 5.13 displays a plot of melting points versus pressure for the sequence 121 (5′-CAT TTA TAA ATG-3′, 16.7% G-C content). Over a pressure range of 0.1-150 MPa, there is a generally linear increase of 1.6 K in the average melting point of the DNA sequence from 313.3 K to 314.9K. The CD melting experiment produced a result of 313.2 ± 2 K at 0.1 MPa. This compares well with the average melting point obtained via NMR spectroscopy of 313.3 ± 2 K at 0.1 MPa.

For residue T11, the proton signals in the spectrum at 316.2 K could not be found. This lack of data meant the melting points for this residue could not be accurately determined and instead had to be estimated (between the ranges of 313.2 K and 319.2 K). Therefore, the estimated melting points were given by the best fit of equation (23). The estimated melting points for this residue are plotted as a dotted line on Figure 5.13 and have been excluded from calculation of the average melting point. There are no melting results reported for 200 MPa, as the data obtained were of insufficient quality to be able to produce reliable results. The melting points for residue T5 were able to be determined for pressures at 0.1 MPa and 50 MPa. However, at pressures of 100 MPa and 150 MPa the shape of the curve made fitting impossible. As such, the melting points at these pressures were visually estimated instead.

Counter to the expectations based on the results presented in **Chapter 4** on hexameric DNA, the melting point of this sequence increased by a small amount with increasing pressure. This is rather unusual considering that sequence 121 is essentially an extension of the sequence used for the DNA hexamer (an A-T sequence flanked by a G-C pairing. This simple increase in the melting point could be attributed to the fact that a longer chain will simply be more stable. However, the change in stability with respect to pressure is more complex compared to the hexamer. The DNA hexamer showed less stability as a result steric forces induced by the small compression which results from the increase in pressure. This combined with only 4 A-T pairings for the rest of the chain has made the hexamer less stable as pressure increased.

If the melting points of each residue of the dodecameric sequence 121 are plotted as a function of position within the sequence (Figure 5.20), one can see clearly how the terminal residues affect the melting point. It can be seen that residues A2 and T11 (inaccurately fitted), those connected to the terminal residues, have a lower melting point than the surrounding residues. This is a sign of the influence on these residues by the terminal residues as described for the DNA hexamer. With the exception of residues T5 and A8, the rest of the residues in the sequence have higher melting points. This increase in melting point indicates an increase in stability of the rest of the sequence. Since the main body of this

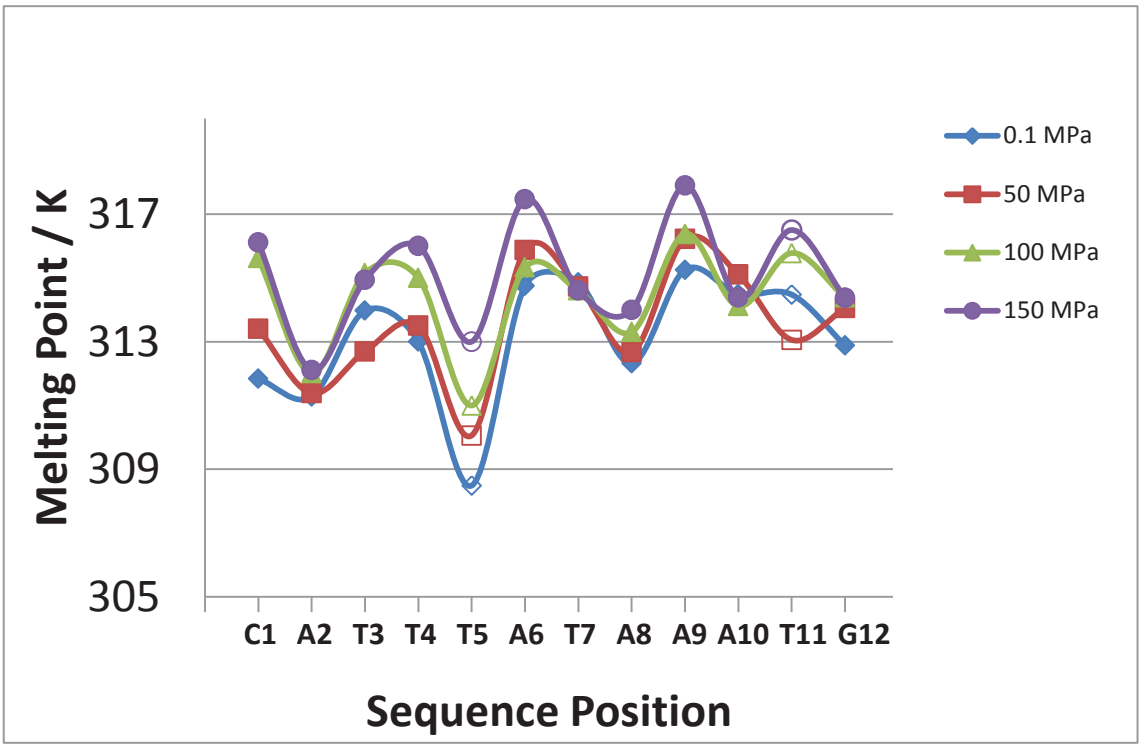


Figure 5.20: Melting Points for Each Residue of Sequence 121 (5′-CAT TTA TAA ATG-3′) Plotted as a Function of Sequence Position. Those melting points which could not be accurately determined are represented by unfilled markers.

sequence is comprised entirely of A-T pairs, there are fewer steric strains on the sequence resulting in pressure-induced stability rather than instability.

Residues T5 and A8 show significantly lower melting points compared to the surrounding residues. This shows that, rather than peeling apart like a zip, the DNA separates from the middle as well as from the ends. Why these residues in particular are more sensitive to pressure is not fully known, but it should be noted that they are a part of two sets of three consecutive residues, with the residues in question being the centre most residues on these sections, TT**T** and **A**AA. This will be discussed in more detail in section 5.4.7.

5.4.2 Sequence 122 (33.3 % G-C Content)

Figure 5.14 displays a plot of melting point versus pressure for sequence 122 (5′-CAT TCT AGA ATG-3′, 33.3% G-C content). Over a pressure range of 0.1-200 MPa, there is a small increase in the average melting point of the DNA sequence from 320.0 K to 321.3 K. Similar to sequence 121, this increase of 1.3 K is somewhat linear in fashion. The CD melting experiment produced a result of 321.2 ± 2 K at 0.1 MPa. This compares well with the average melting point obtained via NMR spectroscopy of 320.3 ± 2 K at 0.1 MPa.

For residue A10 proton signals in the spectrum at 319.2 K could not be found. This lack of data meant the melting point for this residue could not be accurately determined and were instead estimated as with sequence 121. This has resulted in a higher estimated melting point (between the ranges of 325 K and 330 K) with the actual melting point considered to be much lower (within the range of 321 K to 325 K). Additionally, residue T11 also lacked proton signals at 322.2 K and has an estimated melting point in the range of 319.2 K to 325.2 K. The estimated melting points were given by the best fit of equation (23). These estimated melting points have been shown as dotted lines on Figure 5.14 and have been excluded from further examination of the results.

Similar to sequence 121, the melting point of sequence 122 increases linearly by a small amount. This is to be expected since, with the exception of the central four residues, this sequence is identical to sequence 121. Also similar to sequence 121, a plot of melting point as a function of position within the sequence (Figure 5.21) shows that residue A2 has a lower melting point compared to the surrounding bases, the same effect that is seen for the DNA hexamer. As mentioned before, the melting points for residue T11 are estimates and it is considered that they may be a little high. However, it can still be seen that the melting points for T11 are still significantly higher than those within the centre of the sequence.

Two residues in sequence 122 have melting points approximately 12 K lower than the surrounding bases. These are the centre-most T6-6H and A7-8H bases. This decrease in the melting points may have been brought upon by steric strains

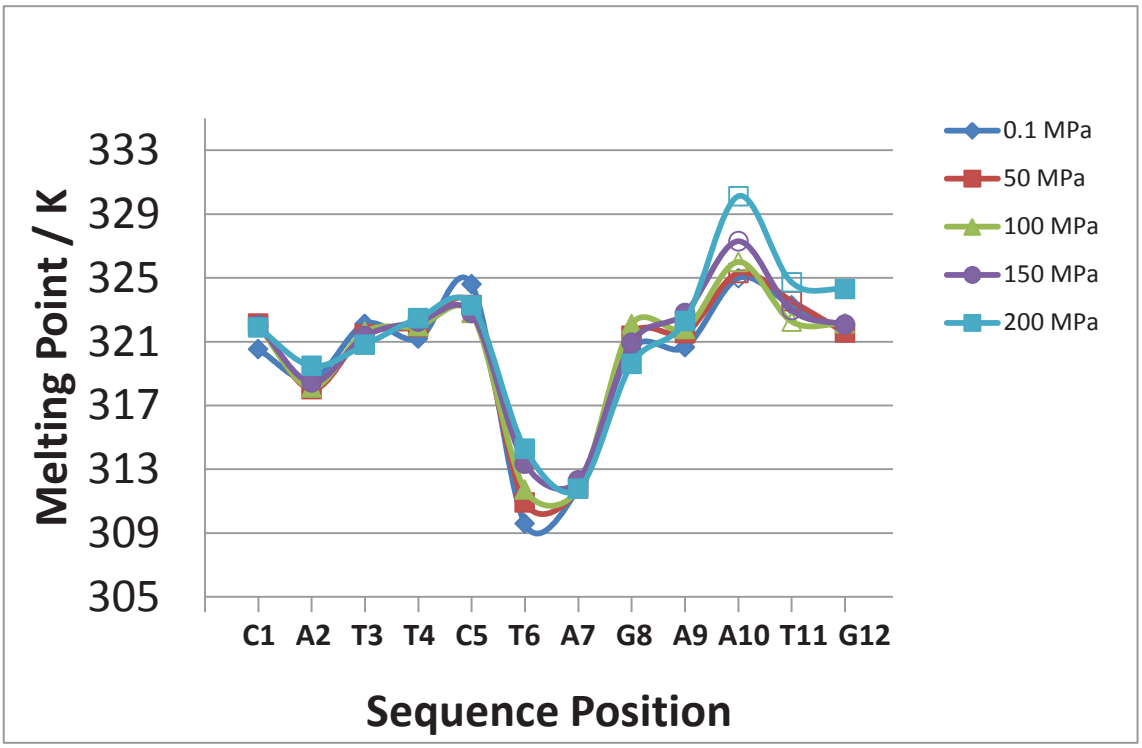


Figure 5.21: Melting Points for Each Base of Sequence 122 (5′-CAT TCT AGA ATG-3′) Plotted as a Function of Sequence Position. Those melting points which could not be accurately determined are represented by unfilled markers.

generated by the flanking G-C pairings (both torsional and bond angle strains). This strain will have made the hydrogen bonding between these T6 and A7 residues less stable, causing then to separate at a lower temperature.

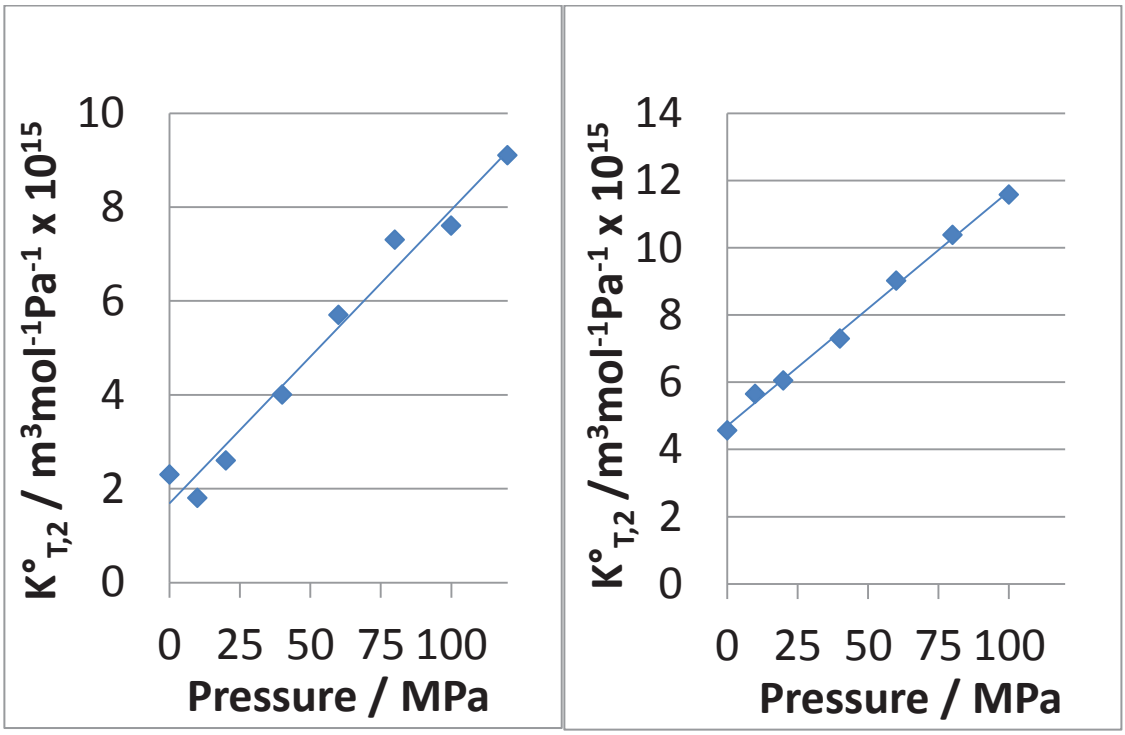
5.4.3 Partial Molar Isothermal Compressions at Infinite Dilution

While both sequences 121 and 122 show increasing melting points under high-pressure conditions, both changes were small and relatively linear. For sequences 123, 124 and 125, a much more interesting change in the melting points with pressure has been observed. Before proceeding to discuss these results, some background information on the effects of pressure on the partial molar isothermal compressions at infinite dilution ($K°_{T,2}$) for individual nucleosides has to be addressed.

Hedwig *et al.* [81-83] have made in-depth studies into the effects of pressure on the partial molar volume of the nucleosides adenosine, cytidine, guanosine and thymidine (and others not relevant to this research). They discovered that each nucleoside exhibited one of two behaviours. In the instances of adenosine and thymidine, $K°_{T,2}$ increases linearly with pressure and is always positive (Figure 5.22). These positive values for $K°_{T,2}$ imply that, over a pressure range of 0.1-100 MPa, the water molecules in the hydration shells of adenosine and thymidine are on average more compressible than the bulk water. There is also no reason to think that this trend will not continue to higher temperatures and pressures.

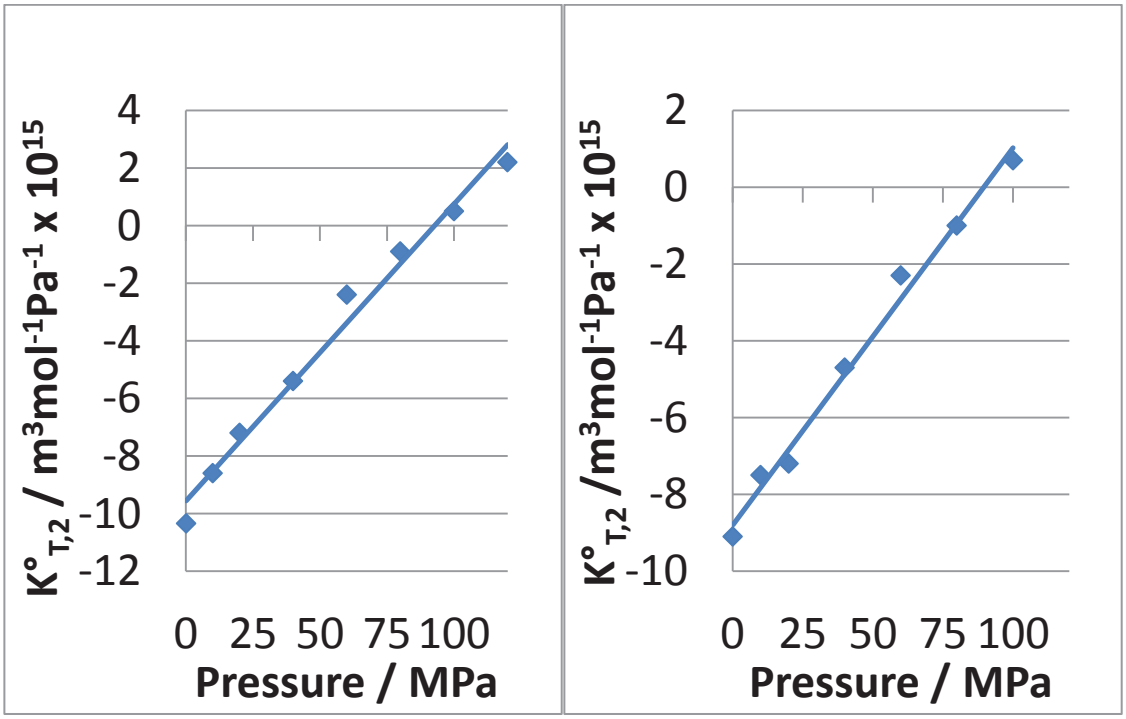
In the instances of cytidine and guanosine, the partial molar isothermal compressions at infinite dilution increase linearly with pressure, starting out negative and becoming positive at approximately 90 MPa (Figure 5.23). The initial negative values of $K°_{T,2}$ imply that the water molecules in the hydration shells of cytidine and guanosine are on average less compressible than the bulk water at pressures lower than 90 MPa. At 90 MPa the compressibility of molecules the hydration shells of cytidine and guanosine are essentially the same as the bulk solvent. At pressures above 90 MPa, the water molecules in the hydration shells of cytidine and guanosine are on average more compressible than the bulk water, much in the same way as adenosine and thymidine.

The consequences of these findings are that water surrounding adenosine and thymidine exhibit opposite behaviours than the water around cytidine and guanosine at pressures below 90 MPa but have similar behaviours above 90 MPa. Comparisons between these results and the results for the pressure dependence of the melting points on the higher G-C content DNA sequences are given below.



Adenosine Thymidine

Figure 5.22: The Partial Molar Isothermal Compressibilities at Infinite Dilution $K°_{T,2}$ versus Pressure for Adenosine and Thymidine at 298.2 K [81, 82].



Cytidine Guanosine

Figure 5.23: The Partial Molar Isothermal Compressibilities at Infinite Dilution $K°_{T,2}$ versus Pressure for Cytidine and Guanosine at 298.2 K [82, 83].

5.4.4 Sequence 123 (50.0 % G-C Content)

Figure 5.15 displays plots of melting point versus pressure for sequence 123 (5′-CAA GTC GAC TTG-3′, 50% G-C content). Over a pressure range of 0.1-200 MPa, there is an overall increase in the average melting point of the DNA sequence from 338.3 K to 342.7 K. However, the average melting temperatures first decreases to a minimum of 337.3 K at 50 MPa before increasing linearly from 100 MPa to 200 MPa. The CD melting experiment produced a result of 338.2 ± 2 K at 0.1 MPa. This compares well with the average melting point obtained via NMR spectroscopy of 338.3 ± 2 K at 0.1 MPa.

For residues C6 and C9, the proton signals in the spectra from 335.2 K to 341.2K could not be found and instead had to be estimated (between the ranges of 332.2 K and 344.2 K). The estimated melting points were given by the best fit of equation (23). These estimated melting points have been shown as dotted lines on Figure 5.15 and have been excluded from further examination of the results.

The fact that the melting point of this higher G-C content DNA sequence initially decreases with pressure before starting to increase after 50-100 MPa draws parallels with the results obtained by Hedwig *et al*. The similarities in the results imply that the compressibility of the water molecules in the hydration shells of the nucleobases plays a large role in the behaviour of the melting point as a function of pressure.

Sequence 123 exhibits a dramatic difference to the results obtained for sequences 121 and 122. Considering the G-C content of the sequences, this difference in results can be explained. The sequences 121 and 122 consist primarily of A-T pairings. This would likely result in their melting trends being governed by the compression behaviour of adenosine and thymidine. Sequence 123, however, is 50 % G-C, and, as such, the compression behaviour of cytidine and guanosine will have a significant effect. The sharp increase in the rate of change of the melting point with pressure can correlated to the larger gradient of $K°_{T,2}$ for cytosine and guanosine compared to those for adenosine and thymidine.

An examination of the melting point of each base as a function of its position within the sequence shows that the terminal bases, C1-6H and G12-8H, and their neighbours, A2-8H and T11-6H, exhibit melting points lower than the rest of the bases in the sequence (Figure 5.24). The central section of the sequence shows a dip in T_m towards the centre when the approximated melting points of C6-6H and C9-6H are ignored.

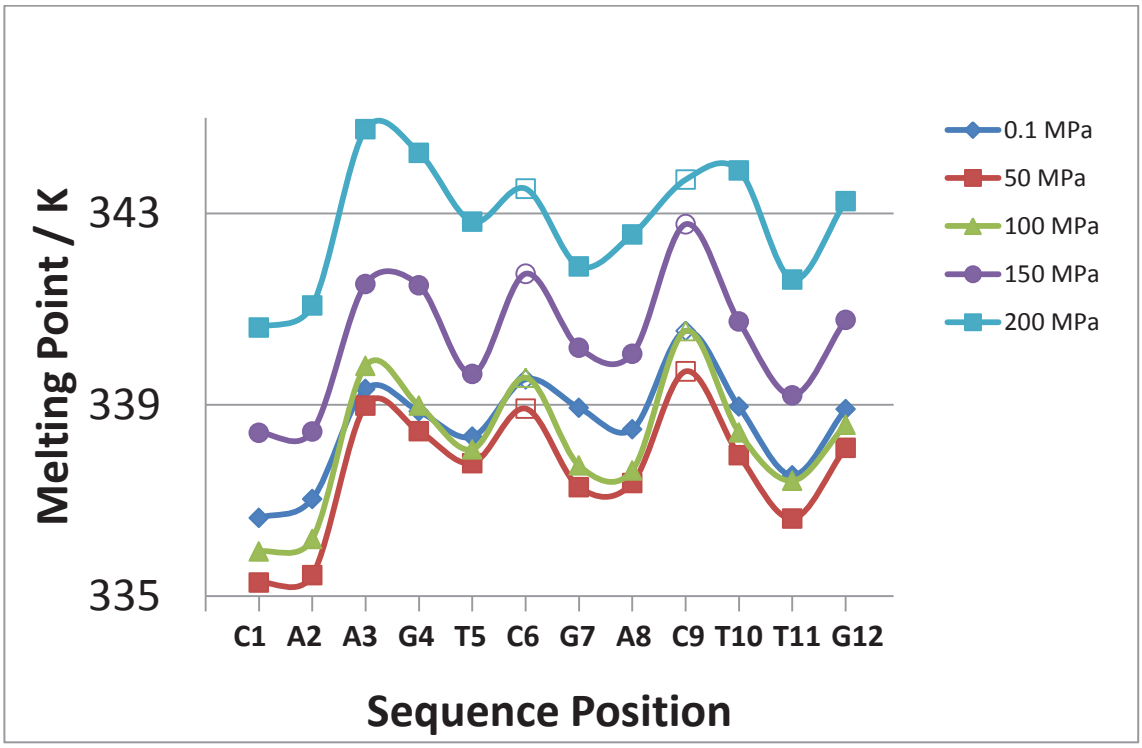


Figure 5.24: Melting Points for Each Base of Sequence 123 (5′-CAA GTC GAC TTG-3′) Plotted as a Function of Sequence Position. Those melting points which could not be accurately determined are represented by unfilled markers.

5.4.5 Sequence 124 (66.7 % G-C Content)

Figure 5.16 displays a plot of melting point versus pressure for sequence 124 (5′-CAG GTC GAC CTG-3′, 66.7% G-C content). Over a pressure range of 0.1-150 MPa, there an overall increase in the average melting point of the DNA sequence from 338.0 K to 339.2 K. Similar to sequence 123, the average melting temperature first decreases to a minimum of 336.3 K at 50 MPa before increasing to the maximum of

339.2 K at 150 MPa. The CD melting experiment produced a result of 338.1 ± 2 K at 0.1 MPa. This compares well with the average melting point obtained via NMR spectroscopy of 338.0 ± 2 K at 0.1 MPa.

For residues C6 and C10, the proton signals in the spectra at 335.2 K, 338.2 K and 341.2K could not be found and instead had to be estimated (between the ranges of 335.2 K and 344.2 K). It should be noted that the estimations made for residue C10 appear to be several degrees too high compared to its complementary base. This is most likely a result of the shape of the curve affecting the fitting process. Additionally, the shape of the melting curve for residue A2 has made it impossible to accurately determine the precise temperature of the melting point and as such, the melting points of this residue have also been estimated in the range of 335 K to 341 K. The estimated melting points were given by the best fit of equation (23). These estimated melting points have been shown as a dotted line on Figure 5.16 and have been excluded from further examination of the results. There are no melting results reported for 200 MPa. The corresponding melting experiment was of insufficient quality to be able to produce reliable results.

The behaviour of the melting point as a function of pressure is very similar to that of sequence 123. This is to be expected as these two sequences are almost identical (A3 and T10 of sequence 123 have been changed to G3 and C10 in sequence 124). An examination of the melting point of each base as a function of its position within the sequence (Figure 5.25) produces results similar to those of sequence 123. The terminal bases, C1-6H and G12-8H, and their neighbours, A2-8H and T11-6H, exhibit melting points lower than the rest of the residues in the sequence and the central section having a dip in T_m towards the centre when the approximated melting points of C6-6H are ignored.

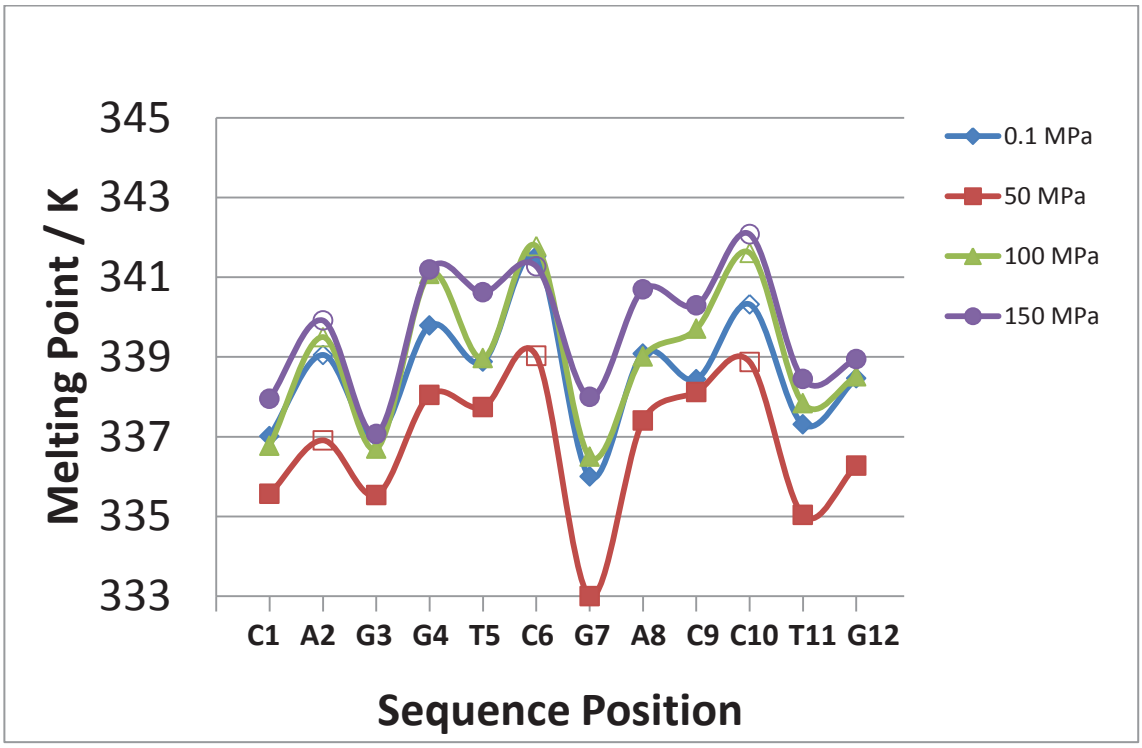


Figure 5.25: Melting Points for Each Base of Sequence 124 (5′-CAG GTC GAC CTG-3′) Plotted as a Function of Sequence Position. Those melting points which could not be accurately determined are represented by unfilled markers.

5.4.6 Sequence 125 (83.3 % G-C Content)

Figure 5.17 displays a plot of melting point versus pressure for sequence 125 (5′-CAC CCG CGG GTG-3′, 83.3% G-C content). Over a pressure range of 0.1-200 MPa, there is an overall increase in the average melting point of the DNA sequence from 346.0 K to 348.7 K. However, the average melting temperature decreases to a minimum of 344.9 K at 50 MPa before increasing linearly from 100 MPa to 200 MPa. The CD melting experiment produced a result of 343.12± 3 K at 0.1 MPa. This result is approximately 3 K lower than the average melting point obtained via NMR spectroscopy of ± 2 K at 0.1 MPa.

For residue C5, the proton signals at 341.2 – 350.2 K could not be found, preventing the accurate determination of the melting point. Instead it had to be estimated (between the ranges of 338.2 K and 353.2 K). It should be noted that the estimations made for residue C5 appear to be several degrees too high. This is most likely a result of the shape of the curve affecting the fitting process.

Additionally, the shape of the melting curve for residue A2 has made it impossible to accurately determine the exact temperature of the melting point and, as such, the melting points of this residue have also been estimated in the range of 341 K to 355 K. The estimated melting points were given by the best fit of equation (23). These estimated melting points have been shown as a dotted line on Figure 5.17 and have been excluded from further examination of the results.

Like sequences 123 and 124, sequence 125 exhibits an initial decrease in the melting point with increasing pressure followed by an increase in melting point after 100 MPa. Since this sequence is comprised almost entirely of G-C base pairs, it is expected that its melting behaviour as a function of pressure will be governed entirely by the pressure-dependent behaviours of cytidine and guanosine. This result further supports notion that the partial molar isothermal compressibilities of the bases in the sequence play an important role in determining the effect of pressure on the melting point of DNA.

In this sequence, an examination of the melting point of each base as a function of its position within the sequence (Figure 5.26) produces results which are slightly different than what was to be expected. The terminal bases, C1-6H and G12-8H, and their neighbours, A2-8H and T11-6H, exhibit melting points slightly lower than the rest of the bases in the sequence. The central section of the sequence shows somewhat variable melting points but these are still relatively constant when considering a level of error of approximately ± 0.5 K. However, the melting point for G8-8H is approximately 6 K lower than its neighbour for all pressures. This was first considered an error but is now considered to be the actual melting point. C5-6H does not show a similar drop, but it was unable to be fitted accurately. This result is the same as that observed for sequence 121, with the residues in question being a part of two sets of three consecutive bases, with the bases being the centre-most bases on these sections, CC**C** and **G**GG.

It should be noted that, once removed from the high-pressure cell, the solution containing sequence 125 had changed from a clear solution to a pale transparent yellow. Examination of the recorded ^{1}H NMR spectra for this sequence yielded no significant unidentifiable signals. Additionally, there is only a small decrease of 5% in the DNA concentration across the time frame during which the sample was

present in the cell. Because of these observations, this change in sample colour is considered to have a zero to minimal effect on the results obtained for this experiment. The cause for this discolouration is unknown and was not present in any other samples, even those which experienced the same temperature and pressure conditions. This degradation of the DNA is possibly the reason the CD spectrum yields a melting temperature 3 K lower than the result obtained from NMR whereas all other CD results compared well to the NMR results at 0.1 MPa.

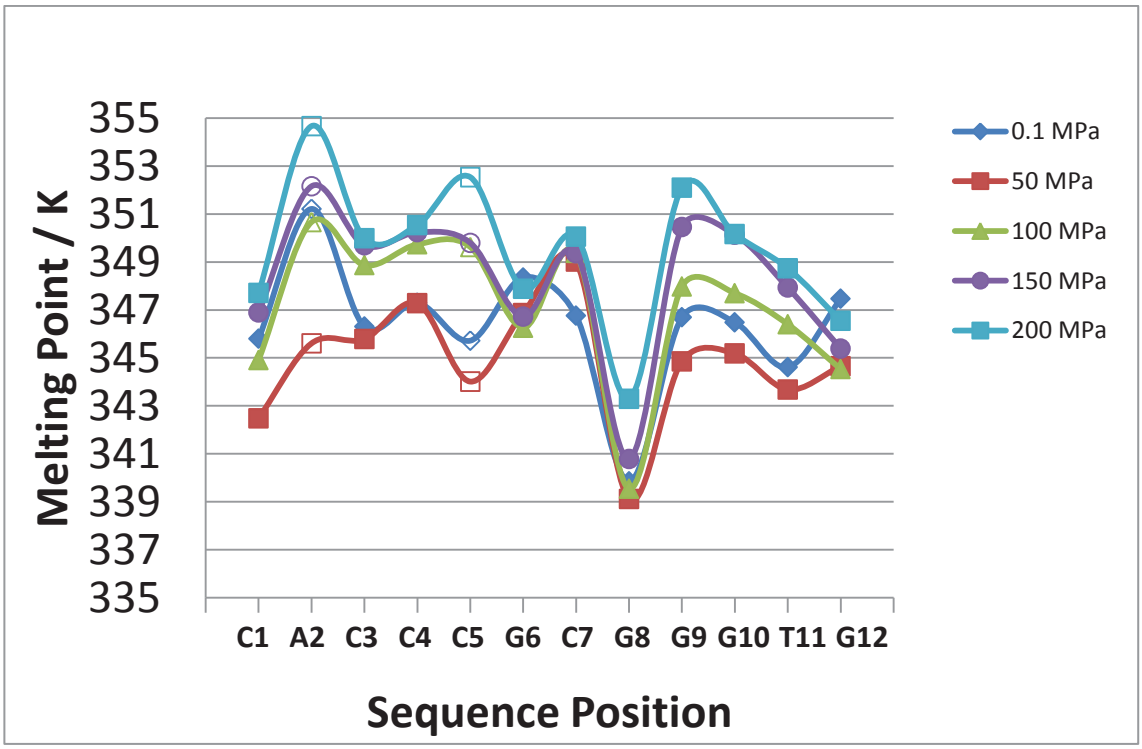


Figure 5.26: Melting Points for Each Base of Sequence 125 (5′-CAC CCG CGG GTG-3′) Plotted as a Function of Sequence Position. Those melting points which could not be accurately determined are represented by unfilled markers.

5.4.7 Influences of Pyrimidines and Purines

The phenomenon of some residues having lower melting points within all the DNA dodecamers is intriguing and in need of explanation. These drops in melting point are believed to be a result of steric forces experienced at junctions between successive sequences of purines and pyrimidines. A visual representation of the

purine and pyrimidine sequences in each DNA sequence, along with an indication of the size of the decrease in the melting points of the bases, is given in Table 5.5. If it is assumed that estimated melting points from residues where NMR signal is lost over a wide temperature range spanning the melting temperature and which do not match the melting points of their complementary bases instead have melting points which do match, a clear trend emerges: that there is a decrease in the melting point of one of the two bases involved in each of the junctions between sequences of multiple purines and pyrimidines. The exact cause of this phenomenon is still unknown and more experiments (see future work) will be required to determine the exact nature of this behaviour. A possibility is that in these instances one of the bases adopts a favourable conformation, transferring all the steric stress to the other base. A second trend is also visible with the drop in melting point increasing with increasing length of consecutive bases. This may

Sequence Position / Sequence	1	2	3	4	5	6	7	8	9	10	11	12
121	C	A	T	T	T	A	T	A	A	A	T	G
T_m Drop											*	
122	C	A	T	T	C	T	A	G	A	A	T	G
T_m Drop										*	**	
123	C	A	A	G	T	C	G	A	C	T	T	G
T_m Drop						**			*			
124	C	A	G	G	T	C	G	A	C	C	T	G
T_m Drop		**				**				*		
125	C	A	C	C	C	G	C	G	G	G	T	G
T_m Drop					**							

Table 5.5: Visual Representation of T_m Drops. T_m drops appear between junctions between consecutive sequences of purines (purple) and pyrimidines (blue). A drop in the melting point of a base is represented by a coloured square below it. Yellow represents a drop of 1-2 K, orange represents a drop of 3-4 K and red represents a drop of > 4 K. A * indicates a melting point which was estimated but agrees with its complementary melting point. A ** indicate an estimated melting point which does not match that of their complementary base. In these instances they have instead been given the T_m drop of the complementary base.

possibly be a result of the melting point of the rest of the sequence increasing rather than the melting point of the stressed sections decreasing. More experiments will be required to determine the exact nature of these melting point drops as the current number of experiments provides rather limited information.

5.5 Conclusions

The results obtained on the effects of high pressures on the thermal stability of DNA dodecamers have shown two different behaviours. For sequences with low G-C content, the melting point of the DNA increases linearly with pressure. However, these increases are small with sequences 121 and 122 having increases of 1.7 K and 0.7 K respectively over the range of 0.1 MPa to 150 MPa. Sequences with higher G-C content have shown an initial dip in melting point between 0-100 MPa followed by an increase in melting point in the range of 100 MPa to 200 MPa. Sequences 123 and 125 showed increases of 4.4 K and 3.0 K respectively over the range of 0.1 MPa to 200 MPa; while sequence 124 showed an increase of 1.1 K over the range of 0.1 MPa to 150 MPa (can be estimated to be an increase of 3.1 K at 200 MPa). The differences in these behaviours have been attributed to the effects of pressure on the partial molar isothermal compressions at infinite dilution, $K°_{T,2}$, of the nucleosides. Both adenosine and thymidine exhibit a continually positive $K°_{T,2}$ value over the ranges of 0 MPa to 120 MPa whereas guanosine and cytosine exhibit a negative $K°_{T,2}$ value between 0 MPa and 90 MPa which becomes positive above 90 MPa.

In addition to information on the effects of pressure on the overall thermal stability of DNA, the behaviour of the individual residues has also been obtained. Some residues have melting points lower than those in the rest of the sequence. The cause of these drops in melting point is believed to be a result of steric strains caused by junctions between sequences of consecutive pyrimidines and purines.

The consequences of these changes in melting point vary depending upon pressure. At pressures below 100 MPa (submarine pressures), the melting point of

DNA increases by a negligibly small amount or decreases (a degree or two depending upon the G-C content). At pressures above 100 MPa, the thermal stability of DNA increases with pressure, with sequences of higher G-C content increasing faster.

These results do not offer much evidence in favour of the submarine high-pressure origin-of-life theory, with pressure having either a negative or only small positive effect on the thermal stability of DNA.

5.6 Future Work

The results obtained from these studies on the effects of pressure on the physical stability of DNA dodecamers have provided some exciting results. However, this has only been an initial study and there are still many experiments which can be performed. These include, longer sequences, examining different sequences and RNA.

5.6.1 Different Sequences

While it has been shown that the G-C content has a significant effect on the melting behaviour as a function of pressure for a DNA sequence, it is still unknown what effect varying the DNA sequence, while keeping the G-C content constant, will have. This may lead to information on how possible steric forces are influenced by the sequence. Additionally, more evidence is needed to support the theory that junctions between pyrimidines and purines result in the observed decreases in melting point. Experiments using longer DNA sequences, including 16-mers and 20-mers, will continue to broaden the understanding of the effects of pressure on the melting point of DNA sequences.

5.6.2 RNA

Repeating these studies using RNA instead of DNA is a must. It has been shown by Hedwig *et al.* [82] that the behaviour of the partial molar isothermal compressions at infinite dilution ($K°_{T,2}$) of uridine is similar to that of cytidine and guanosine. The compressibility of the molecules in the hydration shell of uridine is less than the bulk water below 70 MPa and more compressible above 70 MPa. This cross-over point is a little lower than that of 90 MPA for cytidine and guanosine. This will be expected to lead to RNA sequences with high A-U (instead of A-T) content being more likely to have the same melting with pressure behaviour as those with high G-C content.

Chapter 6 - Physical Stability of i-Motifs

6.1 Introduction

6.1.1 Alternate DNA Structures

Whilst the structures of DNA within living organisms are similar, there is a multitude of different structures which have been observed *in vitro*. Even with the limitation of a double helix, DNA still possesses several different forms [84]: B-DNA (the traditional right-handed DNA helix with bases normal to the helix axis and approximately 10 bases per turn), A-DNA (a right-handed helix with bases inclined around the axis and with 11 bases per turn) and Z-DNA (a left-handed helix formed from alternating G-C sequences which exhibits a distinct "zig zag" backbone and 12 bases per turn). A-, B- and Z-DNA are three of the more well-known DNA structures.

Beyond double helical structures there are also triple-helix structures (H-DNA) and tetrameric associations such as the G-quadruplex (G-DNA) and the i-motif (I-DNA) structures. H-DNA is formed under low-pH conditions, in DNA sequences containing long stretches of both poly-purine (A,G) and poly-pyrimidine (T,C). In these instances, the poly-pyrimidine chain partly dissociates from its complementary strand (Figure 6.1) and folds back parallel to the purine-rich strand of the Watson-Crick duplex. Alternate Hoogsteen [85] base pairings are then formed between the two chains forming the triplex structure.

G-Quadruplex structures are formed by four strands of DNA comprised of entirely of guanine residues [86]. A guanine from each strand forms a G-tetrad by forming Hoogsteen-type hydrogen bonds with two other guanines (Figure 6.2). The resulting G-tetrads then stack, forming a square column of guanines. These quadruplex structures can be formed in three different ways: (i) by four parallel G-

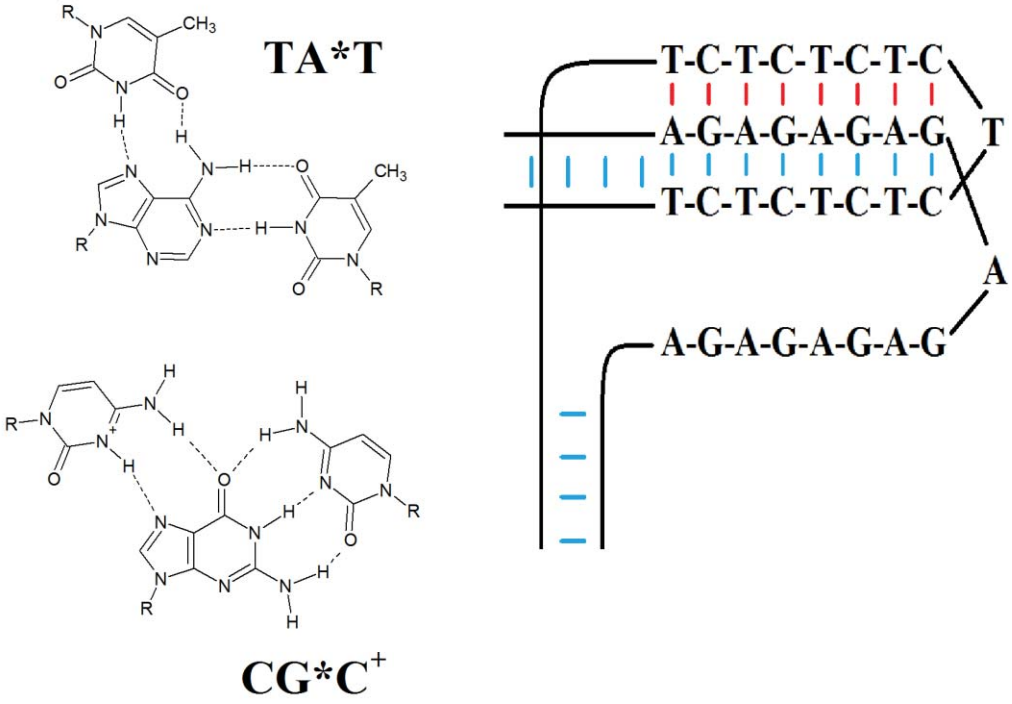


Figure 6.1: Triplex Base Pairing and Folding. a) The T-A*T triplex base pairing. b) The C-G*C⁺ base pairing. These triplex pairings are formed when the central A or G base from both a Watson-Crick base pairing (T-A of C-G respectively) combined with a Hoogsteen base pairing (A*T or C*G respectively). c) The folded triplex structure. A section of the sequence has dissociated from its original duplex pairings and folded back along the chain forming Hoogsteen pairings (red lines) along the Watson-Crick (blue lines) paired duplex.

rich sequences; (ii) from a duplex of two back-folded sequences, each providing two sides of the quadruplex; (iii) and by a single strand back-folding multiple times to form a quadruplex. In addition to the methods of formation, structures involving back-folding can form a multitude of different folded topologies [87] (the formation of which is strongly dependent on the choice of metal ions in the solution [88]). All these possibilities result in a large number of ways in which these G-quadruplexes can be formed.

i-Motif structures are similar to G-quadruplexes in that they are formed by four DNA strands, in this instance comprised entirely of cytosine residues [89] (rather than guanine). However, rather than forming a true quadruplex structure, an i-motif is formed by two antiparallel stranded duplexes connected by C-C⁺ Hoogsteen pairings (Figure 6.3). These pairings are intercalated between the two duplexes resulting in the four-stranded structures. i-Motifs can be formed in the

same way as a G-quadruplex (by four parallel sequences, from a duplex of two back-folded sequences and by a single strand back-folding multiple times). However, there are fewer possible topological structures resulting from back folding because of the antiparallel nature of the two duplexes .

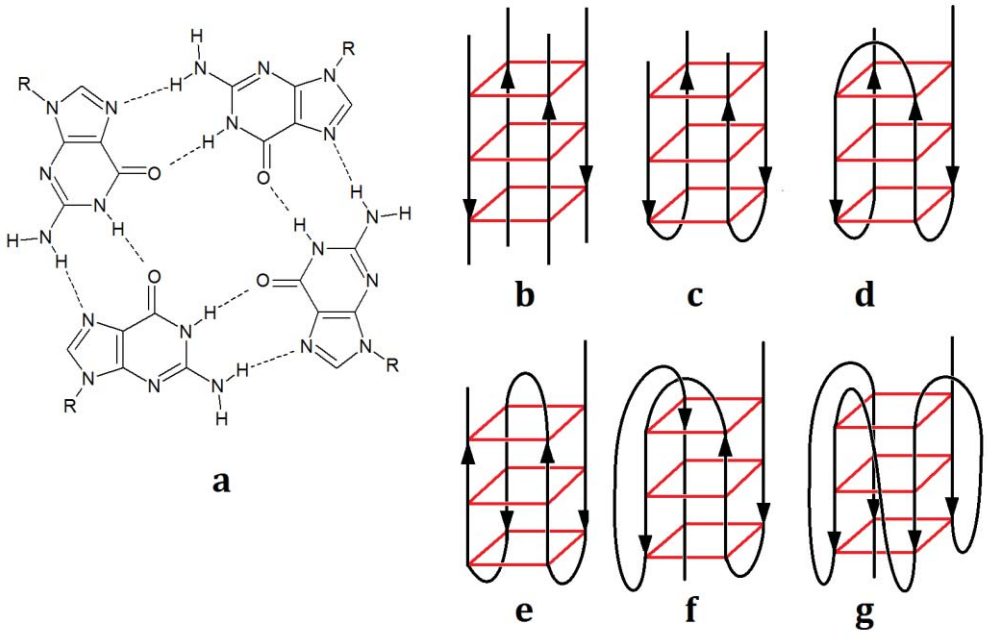


Figure 6.2: G-Quadruplex Base Pairing and Structures. a) The tetrad formed by four guanine residues. b) A G-quadruplex structure formed from four locally continuous oligonucleotides. c) A G-quadruplex structure formed from two separate folded oligonucleotides (one of many different configurations). d) A G-quadruplex structure formed from a single folded oligonucleotide containing three lateral loops. e) A G-quadruplex structure formed from two lateral loops and a diagonal loop, f) A G-quadruplex structure formed from two lateral loops and a propeller loop (the loop around the side of the structure). g) A G-quadruplex structure formed from three propeller loops. The topologies of these structures are only a sample of all the different possibilities.

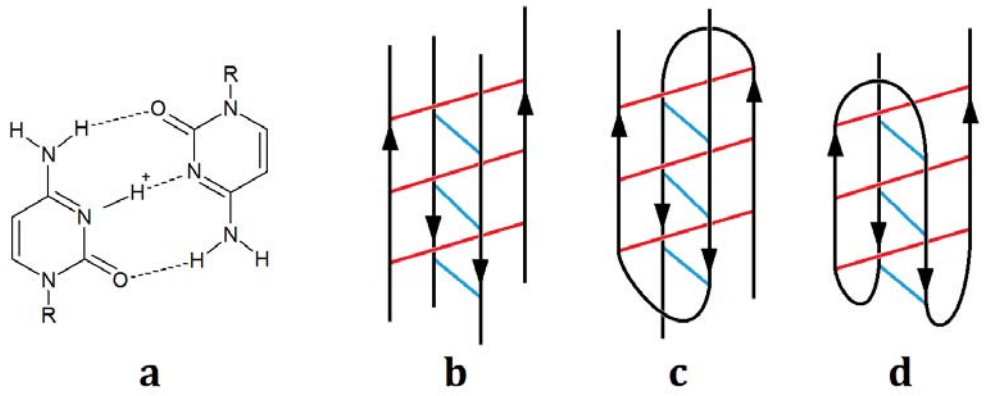


Figure 6.3: i-Motif Base Pairing and Structure. a) the C-C$^+$ base pairing. b) An i-motif structure formed from four separate oligonucleotides. c) An i-motif structure formed from two separate folded oligonucleotides. d) An i-motif structure formed from a single folded oligonucleotide.

6.1.2 Tetramers and Their Importance

Four-stranded DNA structures are of particular interest due to their presence in DNA telomeric sequences. Since their discovery in 1978 [90], telomeres have been known to be important in the cell division process, allowing chromosomes to be replicated without loss of genetic information. DNA replication involves the use of primers attached to the terminal bases at the ends of the DNA sequence. These primers dictate the start and end of the copying process of the DNA but prevent the copying of the sequence where they are attached. This results in the shortening of the DNA sequence upon each replication. Telomeres are, therefore, disposable DNA sequences located at the ends of chromosomes. Each cycle of normal cell division results in the shortening of the DNA sequence without shortening the chromosome itself [91]. In eukaryotes, cells become immortalised when telomeric sequences get regenerated.

Telomeric sequences are comprised of G-rich or C-rich (the complementary strand) repeating units [91]. Some examples (and their complementary sequences) include $(GGG\ TTA)_n$ ($(CCC\ AAT)_n$), found in humans and other vertebrates and slime moulds, and $(GGG\ GTT)_n$ ($(CCC\ CAA)_n$), $(GGG\ GTT\ TT)_n$ ($(CCC\ CAA\ AA)$)$_n$ and , $(AGG\ GTT\ (C/T))_n$ ($(TCC\ CAA\ (G/A)$)$_n$) found in the ciliate protozoa, Tetrahymena and Plasmodium respectively [92]. The length of these repeating sequences varies considerably, from a few hundred bases in ciliates, to thousands of bases in vertebrates [93, 94]. It has been shown that both G-quadruplexes [95, 96] and i-motifs [97, 98] can be formed from these telomeric sequences *in vitro*, with evidence suggesting that this is also possible *in vivo* [99, 100].

6.1.3 Effects of High Pressures on G-Quadruplexes and i-Motifs

There has been very little research conducted on the effects on high pressures on the physical stability of G-quadruplexes, with no known studies performed on the effects on high pressures on the physical stability of i-motifs. Recently Takahashi

and Sugimoto [101] showed that the melting point of the G-quadruplex 15-mer, TBA (5′-GGT TGG TGT GGT TGG-3′), decreases by as much as 14 K at 200 MPa. This effect was seen to decrease considerably in the presence of various crowding agents (ethylene glycol, poly-ethylene glycol M 200 and polyethylene glycol M 4000), reducing the decrease in the melting point at 200 MPa to as little as 5 K. Similar results obtained by Fan *et al.* [102] using the human telomeric sequence 21-mer, 5′-A(GGG TTA)$_3$GGG -3′, also showed that there is a negative effect of pressure on the melting point. Fan's results show that the melting point of their G-quadruplex is reduced by as much as 10 K at 100 MPa (extrapolated to ~20 K at 200 MPa). The differences between these results could be attributed to the differences in the sequences giving different structures, one with two G-tetrads and the other with three G-tetrads.

It should be noted that neither of these studies made any attempt to compensate for any shifts in the pH of the buffer solutions resulting from pressure. The pH 7.00 (pressure 0.1 MPa) Tris/HCl buffer (Takahashi and Sugimoto) and phosphate buffer (Fan) will exhibit, at 200 MPa, pH values of 7.12 and 6.29, respectively. While the changes in the pH of the Tris buffer are insignificant, the changes in the pH of the phosphate buffer are quite large and could also have contributed to the differences in the two observed results.

As stated earlier, there has been no reported research as to the effects of high pressures on the thermal stability of i-motifs. Additional studies along the lines of those already performed on G-quadruplexes are needed to further the understanding of the pressure dependence of more complex DNA structures.

The following is a comprehensive study into the effects of high hydrostatic pressure on the physical stability of the 21-mer i-motif, 5′-CCC TAA CCC TAA CCC TAA CCC-3′. ^{1}H NMR spectroscopy was used to study the effects of high pressures on the melting point of the i-motif under multiple pH conditions ranging from pH 4.1 to pH 6.4. Melting experiments were performed for each pH value at 0.1 MPa and 200 MPa with a small number of additional experiments performed at pressures between these at select pH values.

6.2 Materials and Methods

6.2.1 Sample Preparation

The DNA i-motif, 5′-CCC TAA CCC TAA CCC TAA CCC-3′, was purchased from Integrated DNA Technologies. This DNA, whilst cleaned, still contained impurities from the manufacturing process (particularly the protecting group, benzamide), which are strongly visible in the ^{1}H NMR spectrum. To remove these impurities, the DNA was first purified by precipitation in cold ethanol.

Several 0.20 M citrate buffer solutions were first prepared by adding 0.20 M trisodium citrate stock solution to 2 mL of 0.10 M citric acid stock solution (Table 6.1). The trisodium citrate was added in the presence of a pH meter to ensure the correct pH. Solutions were prepared to pH values as shown in Table 6.2.

Several 0.10 M phosphate buffer solutions were also prepared by adding 0.20 M disodium hydrogen phosphate stock solution to 2 mL of 0.20 M monosodium dihydrogen phosphate stock solution (Table 6.1). The disodium hydrogen phosphate was added in the presence of a pH meter to ensure the correct pH. Solutions were prepared to pH values as shown in Table 6.2.

The DNA was dissolved in 300 μL of milli-Q water and 30 μL of the resulting solution was combined with 150 μL of the desired 0.10 M buffer, 30 μL of 5mM TMSP-d4 stock solution (NMR reference) and 60 μL D_2O (MagniSolv). The solution was made up to 600 μL using milli-Q water, giving final concentrations of: 2 ng μL^{-1} i-motif, 25 mM buffer and 0.25 mM TMSP-d4 in H_2O/D_2O (90:10 v/v). This produced 10 solutions at the pH values given in Table 6.2.

Stock Solution	Mass used /g	Volume used /mL	Concentration /M	Molar Mass /g mol^{-1}	Supplier
$C_6H_8O_7$. H_2O	0.2101	10.00	0.100	210.14	M&B
$Na_3C_6H_5O_7$. $2H_2O$	0.2941	10.00	0.100	294.12	Scharlau
NaH_2PO_4. $2H_2O$	0.3121	10.00	0.100	156.01	Riedel-de Haën
Na_2HPO_4. $2H_2O$	0.3560	10.00	0.100	177.99	Riedel-de Haën
TMSP-d4	0.0073	10.00	0.0050	156.26	Merck

Table 6.1: Table of Stock Solutions for i-Motif Melting Experiments. Each entry shows concentration, mass of compound used, volume made (in milli-Q water) and supplier.

Citrate Buffers	pH 4.1	pH 4.6	pH 5.0	pH 5.5	pH 5.85	pH 6.2
Phosphate Buffers	pH 5.3	pH 5.7	pH 6.0	pH 6.4		

Table 6.2: Buffer pH Values for NMR Spectroscopy Melting Experiments. The pH values for each citrate and phosphate buffer used are shown here.

6.2.2 NMR Spectroscopy

^{1}H NMR spectroscopy along with the high-pressure NMR cell was used for the determination of the melting point of the i-motif under high-pressure conditions.

NMR measurements of the i-motif were conducted with a Bruker Avance 500 MHz NMR spectrometer. Excitation sculpting set at the water offset frequency was used to suppress the dominant water peak in the spectrum.

Atmospheric pressure i-motif melting experiments were performed by recording multiple NMR spectra at different temperatures using a commercial QXI 5 mm ^{1}H-detection inverse probe with a *z*-field gradient coil. NMR spectra were recorded with 3 K temperature steps over a 54 K temperature range using 760 scans and a relaxation time of 2 s, spectral width 24.7 ppm and 64000 points. A time of 3 minutes between each temperature step was allowed for temperature equilibration. These melting experiments were performed at each of the pH values described in the sample preparation.

High-pressure i-motif melting experiments were performed by recording multiple NMR spectra at different temperatures using a commercial TXI 5 mm ^{1}H-detection inverse probe with a *z*-field gradient coil. NMR spectra were recorded with 3 K temperature steps over a 54 K temperature range using 1360 scans, relaxation time of 2 s, spectral width 24.7 ppm and 64000 points. A time of 3 minutes between each step was allowed for temperature equilibration. Melting experiments were performed at 200 MPa for each of the pH values described in the sample preparation with additional melting experiments between the ranges of 0.1 and 200 MPa being performed at pH values of 4.60, 5.50, 5.85 and 6.40.

6.2.3 Circular Dichroism Spectroscopy

Circular dichroism (CD) melting experiments were performed on an Applied Photophysics Chirascan CD spectrometer. A 0.1 mm sample cell was used to reduce the path length through the sample to prevent over saturation without diluting the sample. Melting experiments were performed by recording multiple CD spectra over a temperature range of 278.2 K to 353.2 K in 3 K steps. A temperature equilibration time of 1 minute was used (suitable for the small sample volume).

6.3 Results

6.3.1 Spectral Assignment

The structure of the i-motif is too complex to fully assign the ^{1}H NMR spectrum using the methods outlined in **Chapter 4** and **Chapter 5**. The spectral assignment method will only work for DNA in its B-DNA form, that is, as a right-handed double helix. It is clear that this is not the case with any quadruplex structure. The most that can be achieved is to determine which type of base (A, C, or T) each peak is derived from. Additionally, some assumptions based upon the behaviour of other sequences have been made (Figure 6.4).

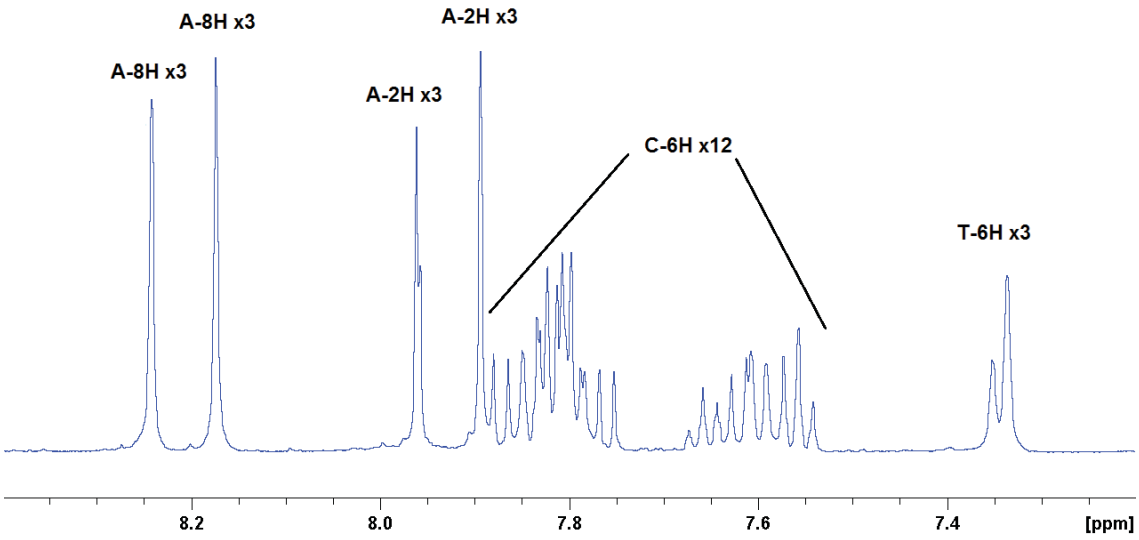


Figure 6.4: i-Motif ^{1}H NMR Spectrum. The types of base for each proton signal have been identified.

6.3.2 Data Gathering

Due to the length and the number of cytosines is this DNA sequence, the resulting ^{1}H NMR spectrum is very convoluted and it is impossible to fully track any individual NMR signals through the melting process. Therefore, no melting curves

could be established as has been done previously. As an alternative, the proton signals corresponding to the H^+ proton (Figure 6.5), involved in the formation of the i-motif, have be monitored instead.

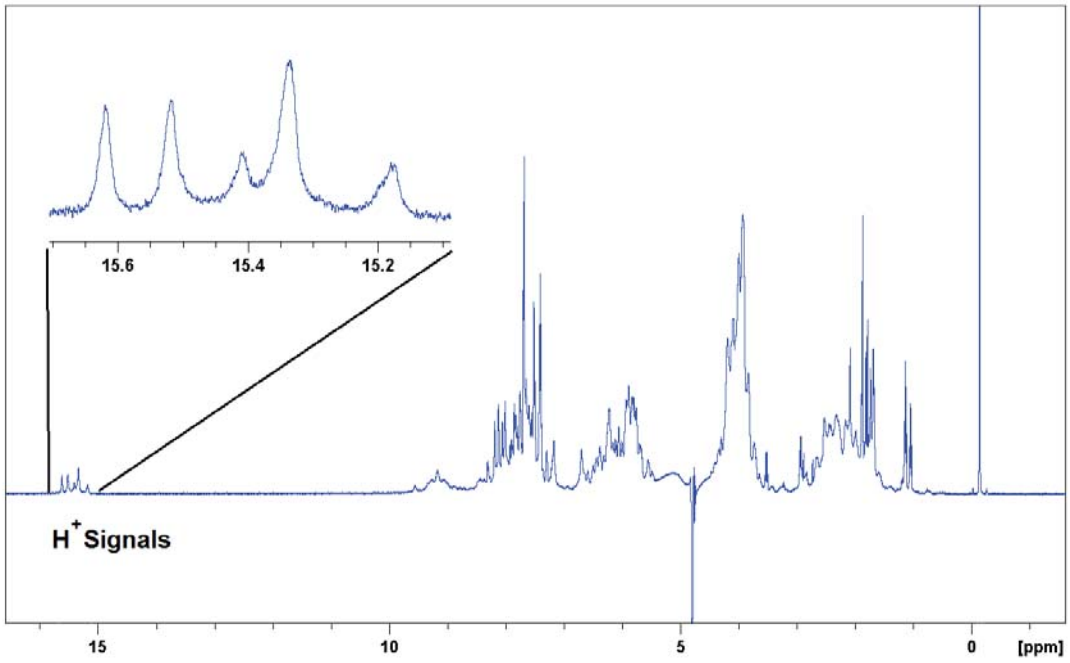


Figure 6.5: i-Motif Spectrum at 285.2 K and pH 6.4. The H^+ proton signal has been magnified to show the fine detail. These signals disappear during the melting of the i-motif.

The intensity of these H^+ proton peaks (TMSP d4) has been monitored as a function of temperature to establish melting curves (Figure 6.6). Intensities of the H^+ peaks were measured by integrating across the whole H^+ signal region relative to the internal NMR standard. By measuring these intensities at each temperature it was possible to establish melting curves which could be fitted using equation (23) to obtain the melting point of the i-motif in solution.

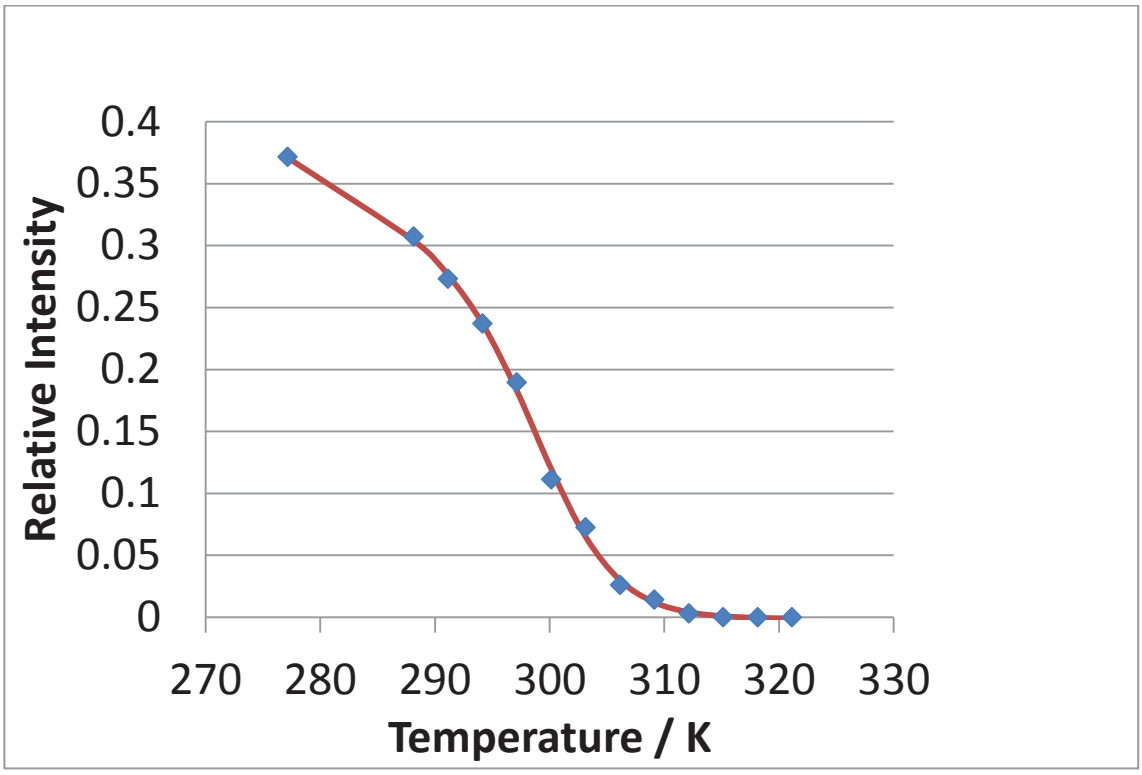


Figure 6.6: Melting Curve for pH 6.4 at 0.1 MPa. The intensity of the H^+ proton peaks relative to the NMR standard has been plotted as a function of temperature along with the fitted melting curve from equation (23).

6.3.3 Effects of pH and Pressure on i-Motif Stability

Figure 6.7 displays a plot of the melting points of the i-motif in phosphate buffer as a function of pH at pressures of 0.1 MPa, 100 MPa and 200 MPa. Values of pH are uncorrected for pressure-induced changes. The pH range over which experiments in phosphate buffer is limited as phosphate buffer has a narrow buffer range and it is not possible to go below pH 4.9. The next buffer range for phosphate is approximately 0.5 to 3.5, too low for these experiments.

Figure 6.8 displays a plot of the melting points of the i-motif in citrate buffer as a function of pH at pressures of 0.1 MPa, 50 MPa, 100 MPa, 150 MPa and 200 MPa. Values of pH are uncorrected for pressure-induced changes

The error in the melting point measurements for both buffers is ± 0.8 K while the error in the pH measurements is ± 0.05.

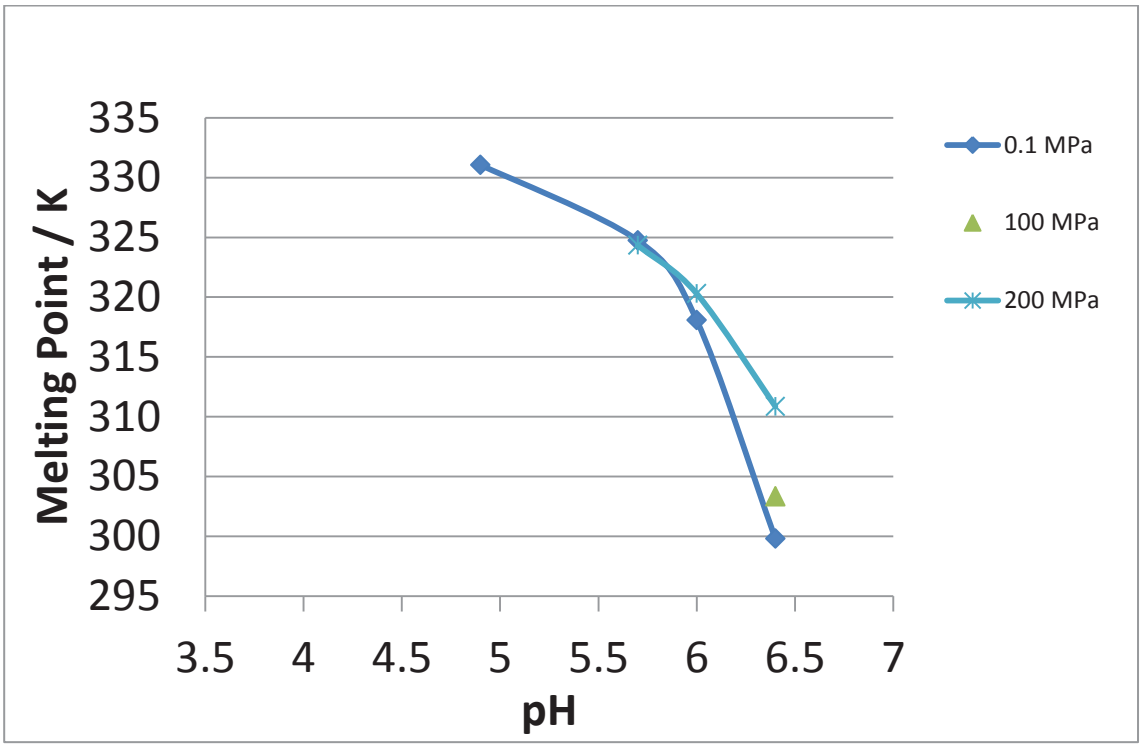


Figure 6.7: i-Motif Melting Points as a Function of Temperature in Phosphate Buffer. Melting points obtained via fitting at pressures of 0.1, 100 and 200 MPa are given. Values of pH are uncorrected for pressure-induced shifts. Lines between data points are included as a visual reference only.

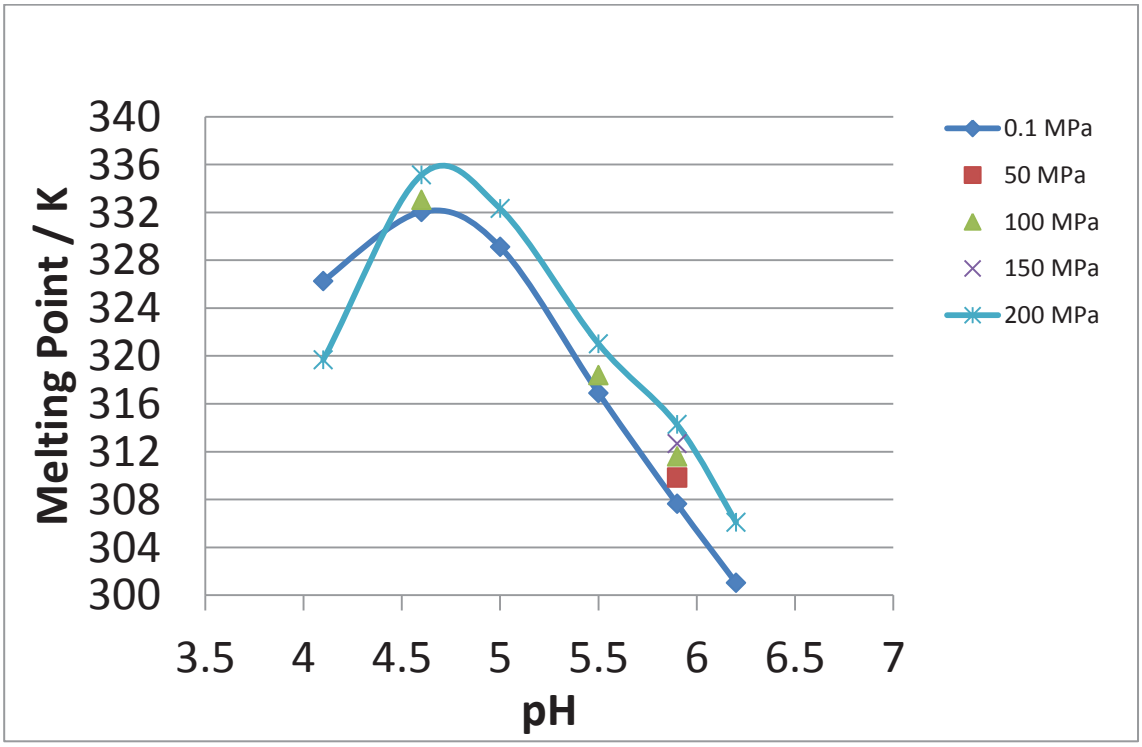


Figure 6.8: i-Motif Melting Points as a Function of Temperature in Citrate Buffer. Melting points obtained via fitting at pressures of 0.1, 50, 100, 150 and 200 MPa are given. Values of pH are uncorrected for pressure-induced shifts. Lines between data points are included as a visual reference only.

6.3.4 Circular Dichroism results

The melting points of a number of samples in citrate buffer were also calculated using CD spectroscopy. The signal intensities of the peaks at 222 nm, 255 nm and 290 nm were recorded, smoothed over 3 nm, normalised and plotted as a function of temperature (Figure 6.9). The melting point of the i-motif sequence is then determined as the average temperature at intensity 0.5. The melting points at 0.1 MPa obtained via CD spectroscopy are presented alongside those obtained by NMR spectroscopy in Figure 6.10. This figure confirms that the melting results obtained via NMR spectroscopy are accurate.

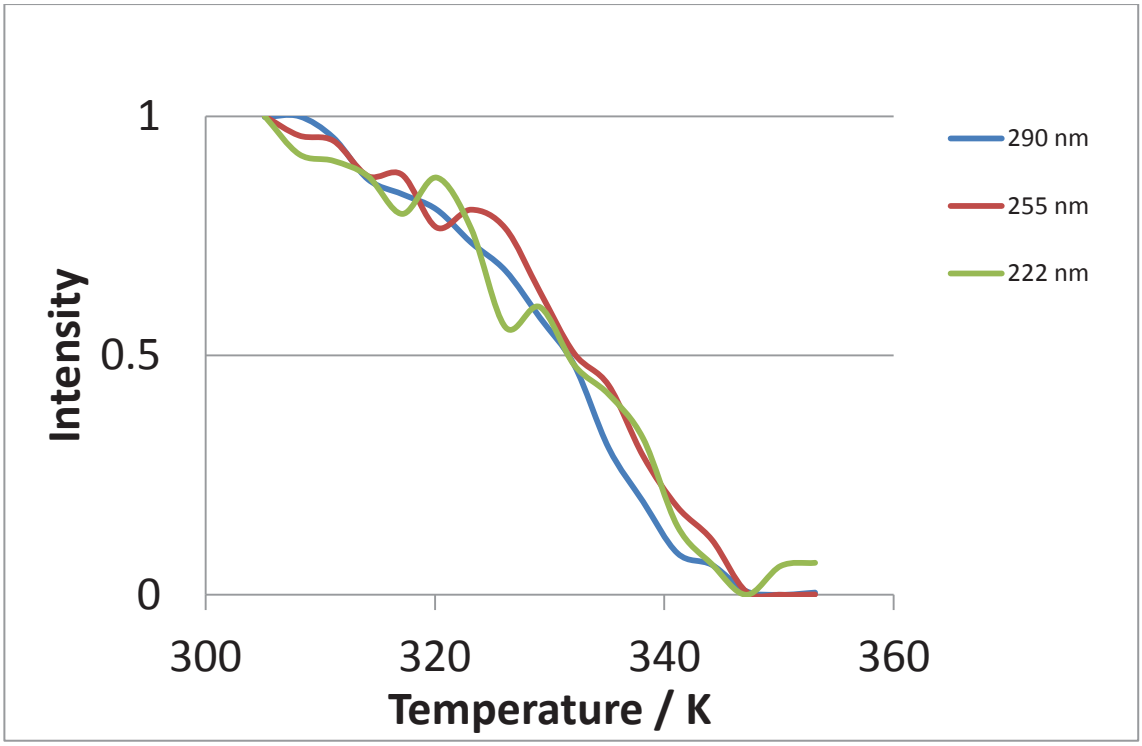


Figure 6.9: CD Melting Curves for the i-Motif at pH 4.6 and 0.1 MPa. Melting curves obtained via CD spectroscopy at wavelengths of 222 nm, 255 nm and 290 nm. The melting point can be determined by the temperature corresponding to the average normalised intensity of 0.5.

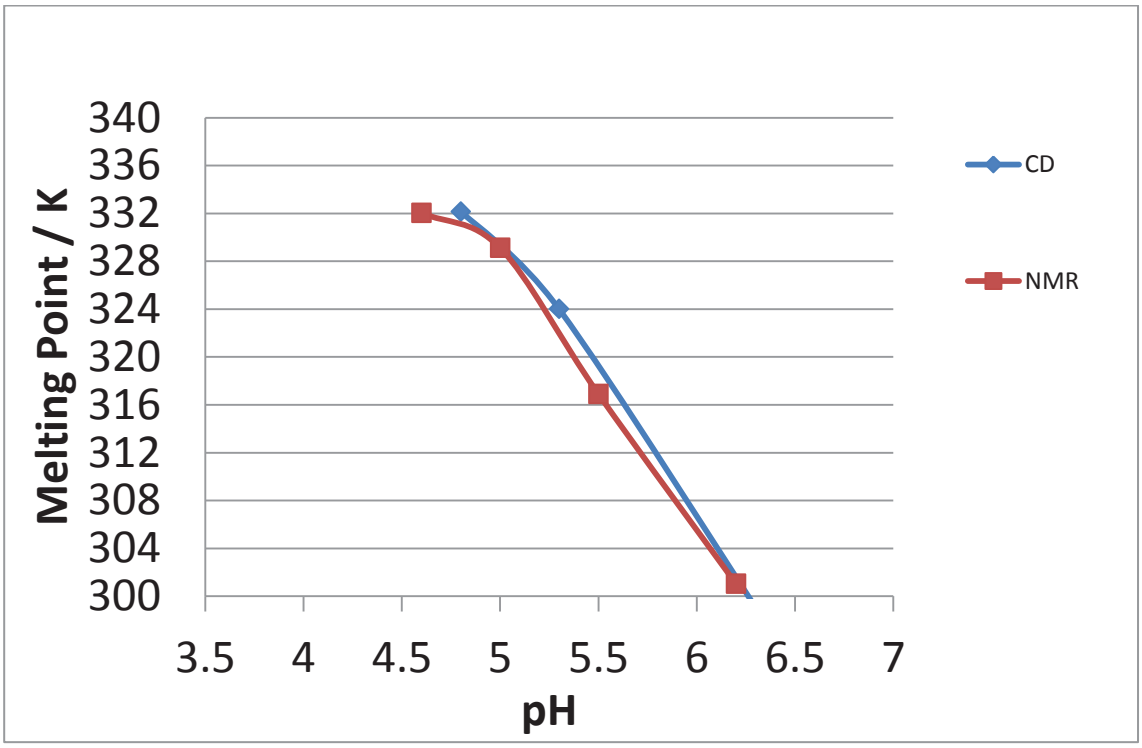


Figure 6.10: Melting Point versus pH for NMR and CD Spectroscopy in Citrate Buffer at 0.1 MPa. Lines are for visual reference only.

6.4 Discussion

6.4.1 i-Motif in Phosphate Buffer

Figure 6.7 shows the melting point of the DNA i-motif, 5′-CCC TAA CCC TAA CCC TAA CCC-3′, as a function of pH at pressures of 0.1, and 200 MPa with an additional data point at 100 MPa at pH 6.4 in phosphate buffer. It is clear in this plot that there is no discernible effect of pressure on the thermal stability of the i-motif at pH 5.7. As the pH increases, the melting point of the i-motif decreases and the effect of pressure becomes more pronounced. At pH 6.4, the difference in the melting points at 0.1 MPa and 200 MPa is 11 K. The melting point at pH 6.4 and 100 MPa is 3.5 K higher than that at 0.1 MPa. This implies that the increase in melting point with pressure is possibly nonlinear in phosphate buffer. However, further measurements will be needed to confirm this.

These results that have been presented so far have yet to consider the effects of the shift in pH induced by the high-pressure conditions. While the pH conditions of an environment may change under high pressures, it is also possible that the pH can be held constant. As such, it is important to have additional analysis of these results with the pH values for the measurements under high pressures corrected, as per section 2.6 to account for their shifts. This is especially important for phosphate buffer which has a ΔV^0 of -24.0 ± 0.4 $cm^3\ mol^{-1}$, resulting in a pH shift of -0.7 at 200 MPa.

Figure 6.11 shows the melting point of the DNA i-motif, as a function of corrected pH, at pressures of 0.1, and 200 MPa. Although the differences in the melting points are less clear, there is no doubt that, unlike before the pH correction, the thermal stability of the i-motif has decreased significantly. Figure 6.12 shows the decrease in the melting point at 200 MPa compared to 0.1 MPa as a function of pH. Readings were obtained by calculating the difference between the melting points at 200 MPa and the 0.1 MPa values at the same corrected pH. The level of the decrease is clearly increasing with increasing pH.

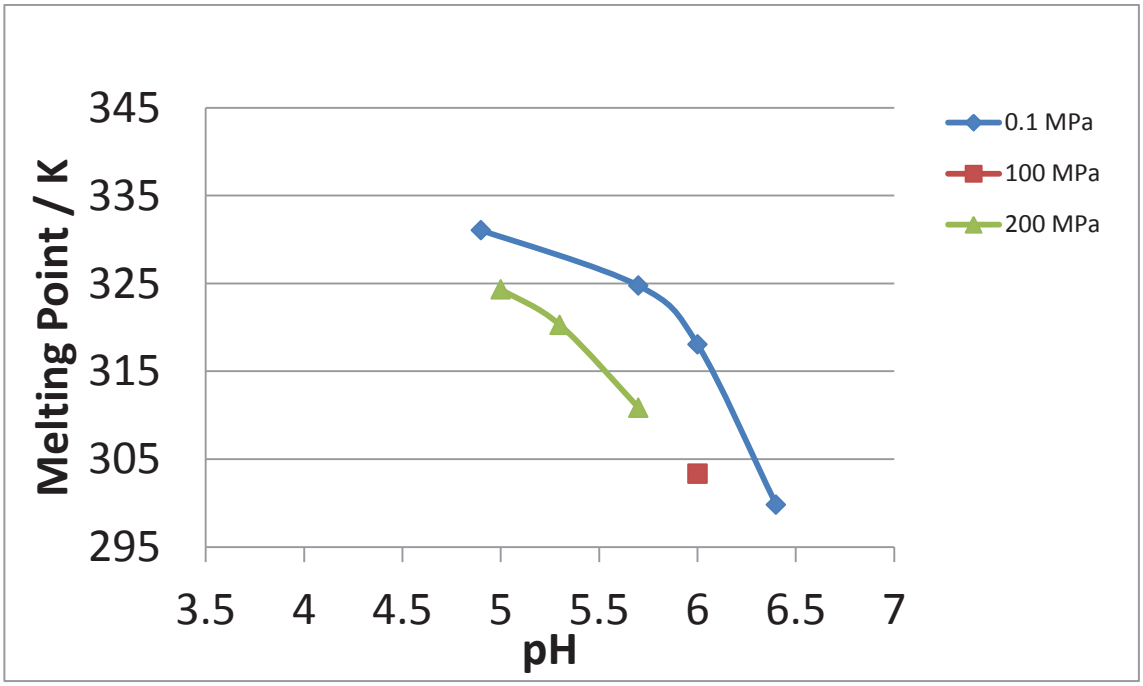


Figure 6.11: i-Motif Melting Points as a Function of Temperature in Phosphate Buffer with Corrected pH Values. Melting points obtained via fitting at pressures of 0.1, 100 and 200 MPa are given with pH values at high pressures corrected for pressure-induced pH shifts. Lines are for visual reference only.

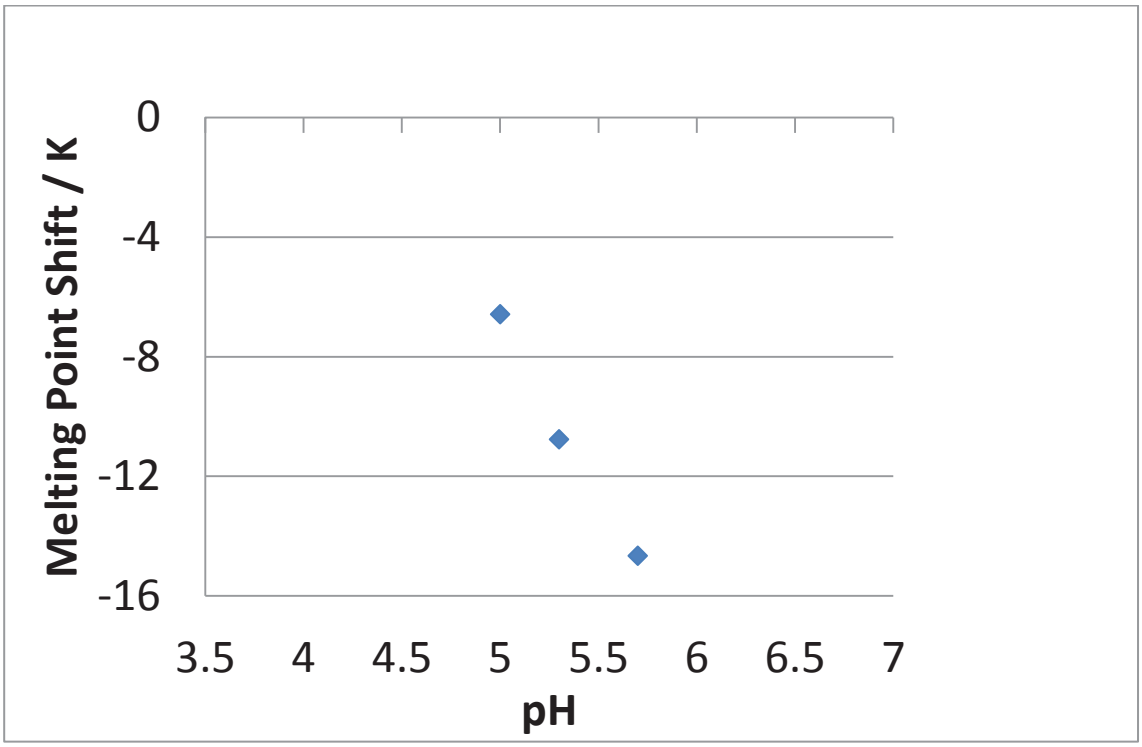


Figure 6.12: Decrease in Melting Point of the i-Motif at 200 MPa. The decrease in the melting point of the i-motif at 200 MPa has been plotted as a function of corrected pH.

6.4.2 i-Motif in Citrate Buffer

As with the phosphate buffer, the effects of pressure on the thermal stability of the i-motif in citrate buffer will be first discussed in the absence of pH correction

Figure 6.8 shows the melting point of the i-motif as a function of pH for pressures of 0.1, 100 and 200 MPa with additional data points at 50 and 150 MPa at pH 5.9. A maximum in the stability of the i-motif at pH 4.6 is observed with the melting point decreasing rapidly with pH either side of this point. This maximum in the stability has been observed before [103, 104] and is dependent on the i-motif sequence. the peak in the melting versus pH curve corresponds well with the pK_a value of 4.6 [105] of nitrogen N3 of cytosine. At this pH there will be a maximum number of C-C$^+$ bonds. At pH values lower than this, protonation of both cytosines occurs while at pH values higher than this there is reduced proton availability for i-motif formation.

Like the phosphate buffer, it is important to assess the high-pressure melting-points of the i-motif accounting for pressure-induced pH shifts. The effect of pH on citrate buffer (ΔV^0 = -12.3 ± 0.4 cm^3 mol^{-1}) is less than that for phosphate buffer; however, there is still a significant pH shift of -0.4 at 200 MPa.

Figure 6.13 shows the melting point of the DNA i-motif in citrate buffer, as a function of corrected pH at pressures of 0.1, 100, and 200 MPa.

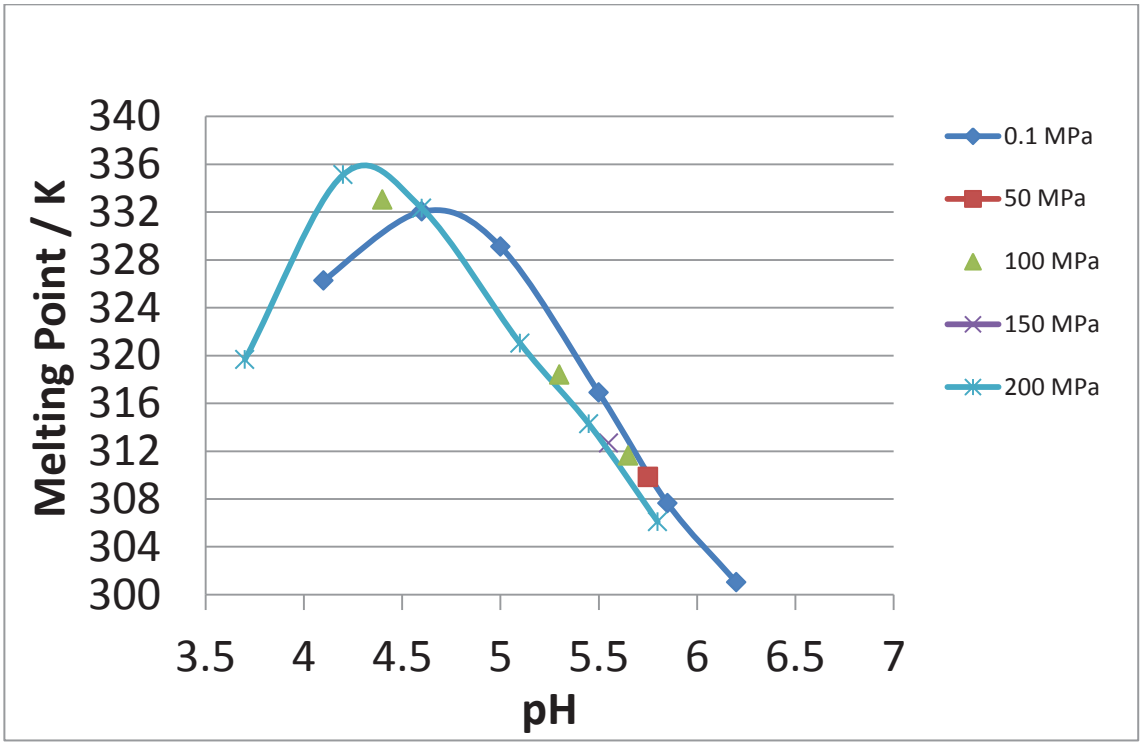


Figure 6.13: i-Motif Melting Points as a Function of Temperature in Citrate Buffer with Corrected pH Values. Melting points obtained via fitting at pressures of 0.1, 50, 100, 150 and 200 MPa are given with pH values at high pressures corrected for pressure-induced pH shifts. Lines are for visual reference only.

Figure 6.14 shows the melting point of the i-motif as a function of pressure at pH 5.85 (no pH corrections). An approximately linear relationship is evident with melting point increasing at a rate of 3.2 K per 100 MPa. When pH changes are considered, it will be true that each different pressure reading in this figure will have a different pH value. At pH 5.9 the melting point versus pressure curve is approximately linear at both 0.1 and 200 MPa. Therefore, at this pH, the relationship between melting point and pressure is pH independent. A second line

on Figure 6.14 shows the change in melting point as a function of pressure relative to the melting point at pH 5.85 and 0.1 MPa. Similar to the uncorrected results, the relationship is approximately linear; however, the gradient is negative with the melting point decreasing by 2.1 K per 100 MPa.

It would appear that an increase in pressure has led to a change in the pK_a of the cytosine 3N nitrogen compared to that at atmospheric pressure. The change in the pK_a of cytosine is not unexpected as it is the same phenomenon which alters the pH of the buffers. The change in pK_a of 0.2 is small but significant. At pH values above 5.0, the relationship between melting point and pH of the i-motif is strongly pH dependant. As a result of this, the small shift in pK_a of 0.2 has resulted in a change in the melting point of 4 K at 200 MPa for the i-motif. At pH values below 5.0, the change in melting point is more complex. The melting-point shift decreases to zero at pH 4.6. At pH values below this, the melting point of the i-motif starts to increase

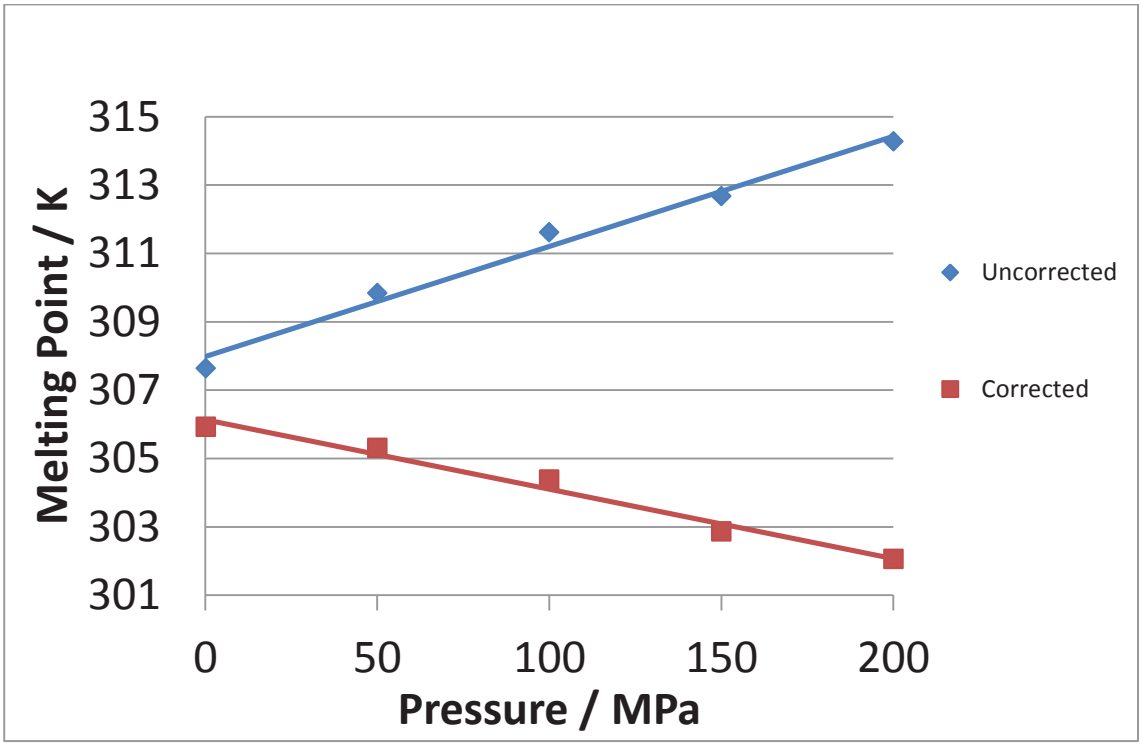


Figure 6.14: i- Motif Melting Point versus Pressure at pH 5.85. The data have had no corrections for pressure-induced pH shifts. The corrected data have been obtained by calculating the decrease in melting point at each pressure compared to the 0.1 MPa value at the same pH. The melting points have then been plotted relative to the pH 5.85 0.1 MPa reading.

with pressure. This is a result of the shifted melting point versus pH curve. The melting point should in theory continue to remain higher under high-pressure conditions. While there is an increase in the melting point with increasing pressure at lower pH values, the biological relevance of this result is lost due to these pH values being far below those seen *in vivo*.

In contrast to this, if pH is not corrected, the change in pH for both citrate and phosphate buffer is significantly large as to increase the melting point of the i-motif by a larger amount than the decrease caused by pressure. This results in an increase of 6 K in the melting point of the i-motif at 200 MPa, rather than a decrease.

6.5 Conclusions

The melting behaviour of the DNA i-motif, 5′-CCC TAA CCC TAA CCC TAA CCC-3′, has been studied at multiple pH values and pressure levels in both phosphate and citrate buffer. In both buffers, the thermal stability of the i-motif increases with increasing pressure in the absence of pressure-induced pH corrections. When the pressure-induced pH shift is accounted for, a decrease in thermal stability is observed for both citrate and phosphate buffer solutions at pH values above 4.6.

These results carry different consequences depending upon the behaviour of pH in natural environments. If pH changes as observed in the laboratory with no correcting factor, high pressures increase the likelihood of i-motif formation. If instead there is some factor which changes the pH of natural environments such that there is no difference in pH under high pressures, then i-motif formation is less likely.

Even given these conclusions, i-motif formation is still largely dependent on the pH of the environment with pressure-induced pH effects playing a larger role on i-motif stability than pressure itself.

6.6 Future Work

The study outlined above has improved the understanding of how high pressures alter the stability of an i-motif structure. However, the solutions used have been simple with very few additional solutes (only buffer and NMR reference), and it is still unknown exactly how other solutes, such as salts and molecular crowding agents, may alter how high pressures affect i-motif stability.

6.6.1 Extra pH Measurements

The current data obtained for the high-pressure melting of the DNA i-motif have been restricted to pH values above 4. It is desired to perform additional experiments at pH values below 4 to fully establish the effects of pressure on the thermal stability of the i-motif over a wider pH range. This can easily be done using citrate buffer; however, for phosphate buffer there is a gap in its buffering range between approximately pH 4 and pH 5. It is not possible to obtain data within this range but it will be possible to obtain data below it.

6.6.2 Salts

As has already been discussed, the presence of salts in DNA solutions has a notable effect on DNA stability and the effect of pressure on stability. Because of this, studies into the effects of the concentration of different salts on the thermal stability of i-motif structures under high-pressure conditions should be performed. The results of this will improve the understanding of how ions may alter the effects of high pressures on i-motif stability.

6.6.3 Molecular Crowding

The studies performed by Takahashi and Sugimoto [101] have shown that molecular crowding has a significant effect on the thermal stability of G-quadruplex structures, reducing the effects of pressure. Given these results, similar experiments should be performed using i-motif structures to see if the same behaviours are observed or whether i-motifs behave differently.

Chapter 7 - Chemical Stability of Cytosine in φX174

7.1 Introduction

After the examining the effects of high temperatures and high pressures on the rate of hydrolysis of cytosine and cytidine as individual bases, the next logical step will be to measure the rate of cytosine hydrolysis within unfolded and folded DNA structures. A paper by Frederico *et al.* describes a sensitive genetic assay used to determine the rate of hydrolysis of cytosine within both the folded and unfolded DNA structures of C141 mutant bacteriophage M13mp2 [28]. This genetic assay, while a rather old method, was remarkably sensitive and capable of observing hydrolysis events as low as 1 in 200,000. They discovered that the rate of cytosine hydrolysis was only marginally faster than for an individual cytosine molecule (8.3 x 10^{-11} s^{-1} for cytosine within the DNA chain compared to 7.9 x 10^{-11} s^{-1} at 303.2 K for individual cytosine). This difference is to be expected as there is little different in the chemical environment. It was also found that within a double-stranded DNA structure, the rate of hydrolysis was 140 times slower than in the single-stranded structure. This is a significant decrease in the rate of hydrolysis and hints at possible higher temperature environments for some of the first life forms. It is important to note that techniques used for this study, though sensitive, are time-consuming and limited in their capabilities. The genetic assay is only capable of recognising, but not distinguishing between, hydrolysis of, for example, C141 and C142 in the C141-C142-C143 codon of the mutant M13mp2 strain. Therefore, the results of this study are only the average rate of hydrolysis of two specific sites.

It was chosen to continue this line of research using ΦX174 [106], a 5386 base circular bacteriophage with 44 % G-C content, and to expand upon it to include studies into the effects of high pressure and high temperature on the chemical stability of cytosine in both single-stranded and double-stranded DNA structures. Next-generation sequencing techniques will be utilised to not only determine the average rate of hydrolysis of cytosine, but also determine the rates of hydrolysis

for each individual cytosine within the entire DNA sequence. This will enable the effects of the neighbouring bases on the hydrolysis rate of cytosine to be examined while simultaneously examining the effects of pressure on these same cytosines.

The following is a comprehensive study into the effects of high pressure on the chemical stability of cytosine within folded and unfolded DNA sequences. DNA samples were incubated at temperatures of 283.2, 313.2 and 333.2 K at pressures of 0.1, 50, 100, and 150 MPa. The incubated samples will then be sequenced using next generation sequencing technology, to identify decayed cytosines which will be observed as a C→T transition. This will enable the determination of the effects of pressure on the rate of cytosine hydrolysis as well as being able to examine whether sequence position has any effects on this behaviour. The planning and implementation of this study has been made possible with the help of Assoc. Prof. Murray Cox.

7.2 Materials and Methods

For these studies, three different forms of ΦX174 bacteriophage were used: single-stranded (SS), super-coiled double-stranded (RFI) and nicked double-stranded (RFII). Samples of these DNA sequences were incubated for up to 2 weeks under the controlled temperature and pressure conditions given in Table 7.1. During this incubation time, samples were removed periodically to establish multiple time points.

The times for incubation were decided upon using the results obtained by Frederico *et al*. An Arrhenius plot of ln k for hydrolysis of cytosine versus 1/T resulted in equation (24)

$$\ln k = \frac{-14494}{T} + 24.1 \qquad (24)$$

where 14,494 is equal to the activation energy (in J) divided by the universal gas constant (E_a/R) and 24.1 is equal to the log of the pre-exponential factor (ln A). This equation can be rearranged to give the reaction constant k (equation (25))

	Incubation Pressure			
Temperature	0.1 MPa	50 MPa	100 MPa	150 MPa
283.2 K	SS (17)[9]			SS (14) [9]
313.2 K	SS (7)[4]	SS (7)[4]	SS (7)[4]	SS (7)[4]
333.2 K	SS (1)[4] RFI (1)[1] RFII (1)[4]	RFII (1)[4]	RFII (1)[4]	SS (1)[4] RFI (1)[1] RFII (1)[4]

Table 7.1: ΦX174 Incubation Conditions. This table outlines the temperature and pressure conditions each form of ΦX174 was subjected to, along with the maximum period of incubation time in days (brackets) and number of time points taken during this time (square brackets).

$$k = exp\left(24.1 - \frac{14494}{T}\right) \tag{25}$$

This can then be inserted into the first-order integrated rate equation to give the fraction of cytosine to uracil transitions, *F*, as a function of time, *t*, and temperature

$$F = 1 - exp\left(t * exp\left(24.1 - \frac{14494}{T}\right)\right) \tag{26}$$

With approximately 50,000 DNA chains being sequenced for each sample and with 1,185 cytosines per chain, it is possible to determine the approximate number of hydrolysis events in 50,000 chains that will occur at atmospheric pressure for each cytosine in the sequence and for each DNA chain. Final incubation times were chosen to give a sufficient number of hydrolysis events at each temperature to obtain the desired information while being short enough to avoid other issues (for example: time constraints and decay of DNA backbone). In addition to this, it has to be considered that the number of events within a double-stranded DNA sequence will be 140 times lower than the single stranded sequences. These incubation times and temperatures, along with the corresponding rate constants, number of hydrolysis events per cytosine in 50,000 strands and the average number of cytosine hydrolysis events in 50,000 chains, are given in Table 7.2 .

Temperature / K	Rate of hydrolysis / s^{-1}	Incubation Time / Days	Number of changes per cytosine in 50000	Average Number of changes per chain in 50000
283.2	2.023×10^{-12}	14	0.1223	145
313.2	2.496×10^{-10}	7	7.548	8944
333.2	3.822×10^{-9}	1	16.51	19564
333.2 (DS)	2.730×10^{-11}	1	0.1179	140

Table 7.2: Incubation Times and Temperatures. Also given are the corresponding rate constants, number of predicted hydrolysis events per cytosine and the average number of predicted cytosine hydrolysis events in 50,000 chains at 0.1 MPa. The last column is for the double-stranded sequence which were incubated at higher temperatures and having a rate constant 140 times lower than its single-stranded counterpart.

7.2.1 Sample Preparation

All water used in sample preparations was milli-Q water, and masses and volumes used in preparation of all buffer stock solutions are given in Table 7.4.

Stock Solution	Mass used /g	Volume used /mL	Concentration /M	Molar Mass /g mol^{-1}	Supplier
$NaH_2PO_4.2H_2O$	1.5601	50	0.2	156.01	Rieded-de Haën
$Na_2HPO_4.2H_2O$	1.7799	50	0.2	177.99	Rieded-de Haën

Table 7.3: Table of Buffer Stock Solutions for Chemical Stability Experiments. Each entry shows concentration, mass of compound used, volume made (in milli-Q water) and supplier.

Solutions of ΦX174 bacteriophage with concentrations of 1000 ng μL^{-1}, in single-stranded, double-stranded super-coiled and double-stranded nicked forms, were

purchased from New England Biolabs. Each DNA solution was diluted 10 fold to 100 ng μL^{-1} solutions.

Several 0.20 M phosphate buffer solutions were also prepared by adding 0.20 M disodium hydrogen phosphate stock solution to 0.20 M monosodium dihydrogen phosphate stock solution. The disodium hydrogen phosphate was added in the presence of a pH meter to ensure the correct pH. The pH values of each buffer solution are such that they will be at the desired pH when under high-pressure conditions as given in Table 7.4.

For each incubation condition, 100 µL of DNA solution was prepared (Table 7.1). The incubation solutions consisted of 20 % V/V 100 ng μL^{-1} DNA solution, 25 % V/V 0.1 M phosphate buffer (for the given pressure) and 55 % V/V milli-Q water. These solutions were prepared in advance and stored in Eppendorf tubes in a -20°C freezer.

In preparation for incubation, the solutions were thawed and 50 µL aliquots inserted into two 75 µL capillaries for each sample to be incubated (this is to provide a backup for each sample). The capillaries were capped with paraffin oil and inserted into the appropriate incubation vessel.

Pressure / MPa	pH under pressure	pH shift	Corrected pH at 0.1 MPa
0.1	7.00	0	7.00
50	7.00	0.20	7.20
100	7.00	0.39	7.39
150	7.00	0.56	7.56
200	7.00	0.71	7.71

Table 7.4: Buffer pH Corrections. This table displays the buffer solutions used for these studies. The first two columns show the intended pressure conditions and the pH required. The third column shows the resulting pH shift and the fourth column shows the pH value at 0.1 MPa required to achieve the correct pH under pressure.

7.2.2 Incubation

Samples to be incubated at pressures greater than 0.1 MPa were inserted into the capillary holder of the high-pressure reaction vessel described in **Chapter 2**. The reaction vessel was filled with paraffin oil and pressurised to the desired level using the high-pressure pump system.

Samples to be incubated at 0.1 MPa where were placed in a small steel vessel filled with paraffin oil. The steel vessel filled with oil replicated the physical conditions within the high-pressure reaction vessel and allowed for parallel incubations at different pressures to be run.

Before samples were inserted, the reaction vessels were equilibrated at the desired incubation temperature for at least 12 hours.

For incubations at 313.2 K and 333.2 K, vessels were placed in a large 80 L water bath. For incubations at 283.2 K, vessels were placed in a temperature-controlled cold room.

7.2.3 PCR Amplification

During the incubation process, a number of cytosines undergo hydrolysis to uracil. This presents a problem during the sequencing procedure. An important step in the sequencing process requires the amplification of the DNA stands involved. However, standard polymerases are unable to copy uracil. To solve this problem, an initial PCR amplification of each DNA sample was performed before it is sequenced. To achieve this, a uracil-tolerant polymerase (KAPA HiFi Uracil+ ready mix purchased from Kapa Biosystems) was used. This polymerase works by copying each uracil as a thymine resulting in a net C-T transition for each hydrolysed cytosine. The KAPA HiFi Uracil+ ready mix contains 0.2 mM of each dNTP and does not contain dUTP.

Template DNA was prepared by addition of 2 μL of each incubated DNA sample to 38 μL of milli-Q water. This gave template concentrations in the range of range of 1 - 0.3 ng μL^{-1}.

PCR amplification was performed by Dr Mark Patchett. The PCR mix (Table 7.5) was prepared according to the manufacturer's specifications and the PCR protocol (Table 7.6) was optimised using the manufacturer's instructions as a starting point.

After PCR amplification, the resulting products were cleaned with a commercial PCR clean-up kit. This produced a 30 μL solution of each amplified sample in a 10 mM pH 8.5 Tris/HCl buffer. The quality of each sample was then confirmed by gel electrophoresis and concentration determined using a nanodrop spectrophotometer.

Component	**Initial concentration**	**Volume added / μL**	**Final concentration**
2 X KAPA HiFi Uracil+	2X	20	1X
Template	1 - 0.3 ng μL^{-1}	1	0.025 - 0.0075 ng μL^{-1}
Forward Primer (IDT) 5'- CCC AAT GCT TTG CGT GAC TAT-3'	5 μM	1.5	0.1875 μM
Reverse Primer (IDT) 5'- CGA ACG TCA GAA GCA GCC TTA-3'	5 μM	1.5	0.1875 μM

Table 7.5: PCR Mix. This table shows each component added to the PCR mixture. Each component's initial concentration, volume added and final concentration are given. The final solution was made up to 40 μL using milli-Q water.

Step	Temperature /°C	Time /s	Cycles
Initial Denaturation	95	180	1
Denaturation	97	25	20
Annealing	60	20	
Extension	72	600	
Final Extension	72	600	1

Table 7.6: PCR Protocol. Temperatures, times and cycles used during PCR amplification of the incubated DNA samples.

7.2.4 Sample DNA Integrity Checks

After completion of incubation, gel electrophoresis was used to examine the length of the DNA to ensure that an unacceptable level of DNA degradation has not occurred. This was performed using a 0.8 % agarose gel using a 1 kb plus ladder (Invitrogen) as a reference. Gels were run for the longest incubation time for each incubation condition to insure sufficient intact DNA remained for PCR amplification. After PCR amplification, a second gel was run for each sample to insure that a sample of sufficient quality was provided for sequencing.

7.3 Next-Generation Sequencing

The next step after amplification of the incubated DNA samples is to determine the level of cytosine hydrolysis. The original work performed by Frederrico *et al*. used a novel genetic assay which would return a positive test in the event of the hydrolysis of one of two specific cytosines in the sequence studied. This method, while successful, is generally outdated and only produces information about two cytosine sites in the sequence. A goal of this research is to gain a large quantity of information on the effects of pressure on the hydrolysis of cytosine at individual sites within the DNA sequence. This will require parallel sequencing of the DNA on

a large scale (of the order of 50,000 strands) for each DNA sample obtained from incubation. This will require next-generation DNA sequencing techniques. The chosen method was Illumina's sequence-by-synthesis (SBS) technology. This method allows whole genome sequencing and supports massively parallel sequencing making it ideal for the desired outcome.

7.3.1 Sequencing by Synthesis

SBS technology works by copying the DNA strand to be sequenced using modified deoxyribonucleotides (dNTPs) [107]. These dNTPs contain a terminator group which blocks further polymerization. This restricts the rate of polymerisation so only a single base can be added by a polymerase enzyme to each growing DNA copy strand. The terminating group for each dNTP possesses a different coloured fluorescent marker which can be detected by a camera to identify which base has been added to each strand. Since all four dNTPs are present during each sequencing cycle, natural competition minimizes incorporation bias. This sequencing reaction is conducted on a large scale across millions of different template molecules laid out on a solid surface. This allows the determination of the next base in the sequence to happen simultaneously across all the template molecules using a signal camera image (Figure 7.1). The fluorescent terminating groups are enzymatically cleaved to allow for the addition of the next base in the sequence. This process is repeated to establish the full DNA sequence.

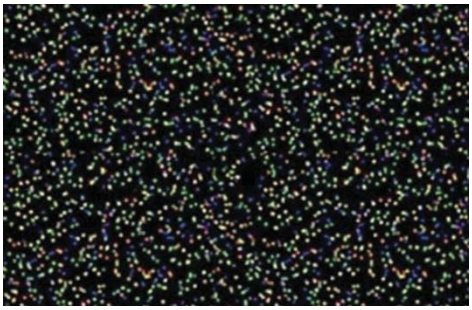


Figure 7.1: SBS Sequencing Image. Each of the differently coloured dots in this image corresponds to a given DNA base. Computer software will examine each sequencing image to construct the sequence for all the DNA strands on this surface. Image obtained from Yale Centre for Genome Analysis.

7.3.2 The Illumina Process

To allow for successful sequencing using SBS, several important steps are required for preparation of the DNA [108]. A general outline of this process is given below along with a visual representation given in Figure 7.2.

The first step involves random shearing of the full DNA sequence (template) to 200-300 base-pair fragments. This is done with either mechanical or enzymatic shearing. This fragmenting of the main sequence will allow faster sequencing times as multiple sections of the chain can be sequenced in parallel.

A sequencing library is then prepared by attaching custom adaptors to each end of the template fragments. Each adaptor corresponds to a different sample which is to be sequenced, allowing multiple samples to be run simultaneously. The library is passed through a flow cell where the adaptors bind the template fragments to wells on its surface.

The fluorescent imaging technology used in Illumina sequences lacks sufficient sensitivity to detect fluorescence signals from signal fluorescent molecules. To improve the signal in order to enable detection, solid-phase “bridge amplification” PCR is used to multiply the strands attached to the flow cell. This generates tight physical clusters containing approximately one million copies of each template fragment. These clusters provide sufficient fluorescent signal to be detected.

The template fragments are then sequenced as described above. This produces the sequences for each template fragment. These sequence fragments are then assembled using a computer algorithm which identifies overlapping sections in the fragments, which allows them to be pieced together.

Illumina sequencing of the incubated ΦX174 bacteriophage samples was performed by New Zealand Genomics Limited (NZGL). All sample processing beyond PCR amplification (enzymatic shearing and library preparation) was also conducted by NZGL.

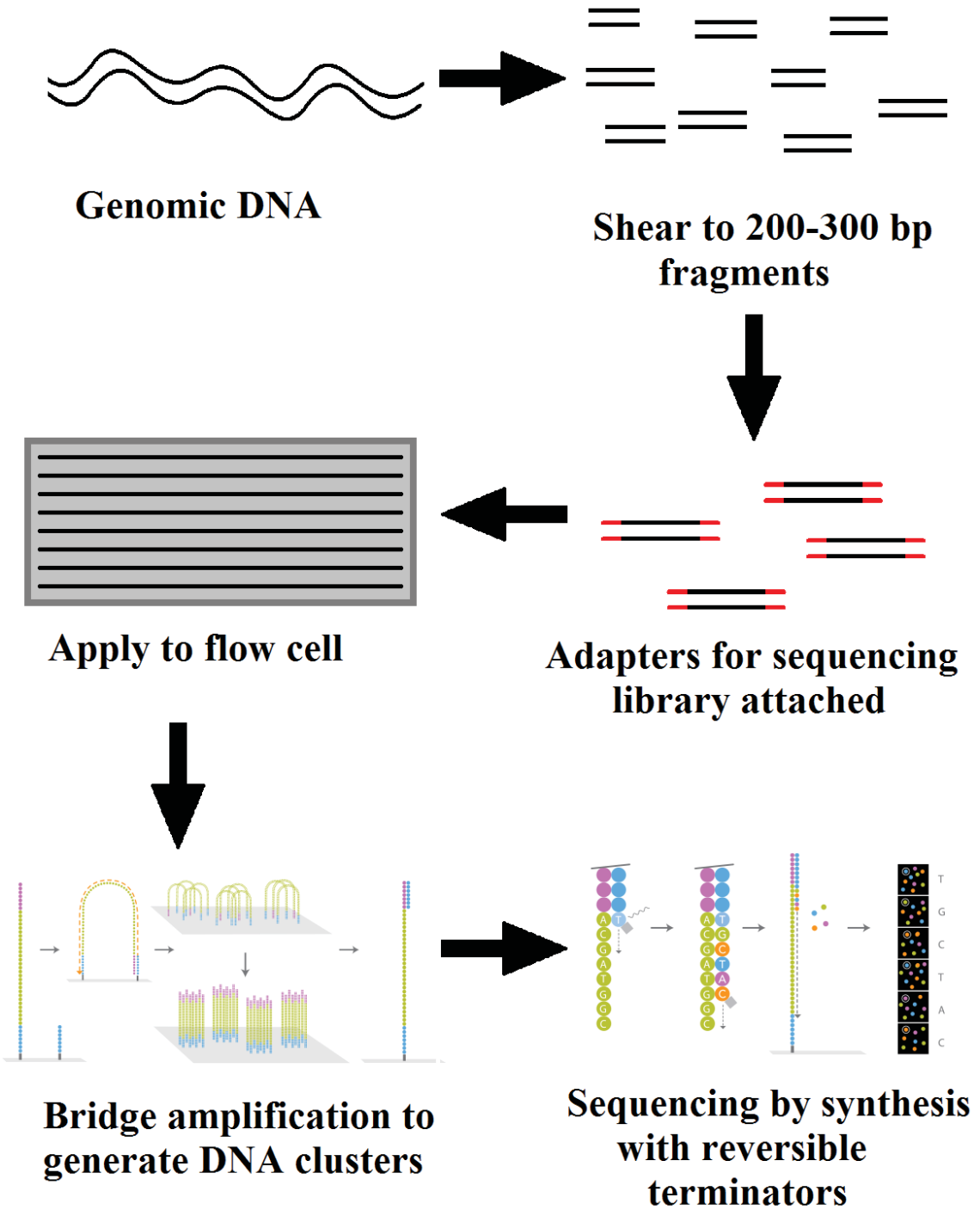


Figure 7.2: Illumina Preparation and Sequencing. Genomic DNA is first sheared to lengths of 200-300 bp. Adapters are added to the ends of the DNA fragments and then flowed through a flow cell where they bind to designated sites. Bridge amplification is then used to generate DNA clusters in preparation for sequencing by synthesis. Image adapted from Bite Sized Bio, http://bitesizebio.com/13546/ sequencing-by-synthesis-explaining-the-illumina-sequencing-technology/.

7.4 Results

7.4.1 Pre Amplification Gel Electrophoresis

After the incubation process, the longest incubated sample was assessed with gel electrophoresis to ensure the quality of the DNA. The following are comments regarding the quality of the samples from a range of different temperature and time conditions which were considered. These results were used in determining the final temperature and time conditions used for this experiment.

- 283.2 K, 2 weeks. Lanes 2 and 3 in Figure 7.3 correspond to these conditions. Lane 2 corresponds to a sample that had been incubated at atmospheric pressure and had very little signal. A second gel analysis of this sample was performed which resulted in a strong band. This difference can be attributed to poor mixing of the sample after thawing. Lane 3 corresponds to a sample incubated at a pressure of 150 MPa and displays a strong band. It was decided that any lower than this temperature would be too close to any base-line hydrolysis which would have happened during sample preparation.

- 313.2 K, 1 week. Lanes 4, 5, 6, and 7 in Figure 7.3 correspond to samples subjected to these conditions. With the exception of lane 4, all these lanes produced bands which are strong and clear indicating good quality DNA. Lane 4 corresponds to a sample incubated at atmospheric pressure and shows very little signal. A second gel analysis of this sample gave a strong band. This difference can be attributed to poor mixing of the sample after thawing.

- 353.2 K, 1 week. Under these conditions a white deposit was seen to develop on the inside of the sample capillaries at all pressures. Subsequent gel electrophoreses confirmed that this was a result of DNA degradation.

- 353.2 K, 1 day. With the reduced time frame there was no clear sign of the white deposit on the tubes of these samples. However, gel electrophoreses

showed no clear bands for the incubated samples indicating a bad quality sample.

- 333.2 K, 1 day. The remaining lanes in Figure 7.3 (8-15) correspond to samples subjected to these conditions. All bands with the exception of those in lanes 8 and 14 are clear and strong. Lane 14 is a little weak but still acceptable. Lane 8 corresponds to a sample incubated at atmospheric pressure and shows a very weak band. A second gel analysis of this sample was performed which showed a strong band. This difference was attributed to poor mixing of the sample after thawing.

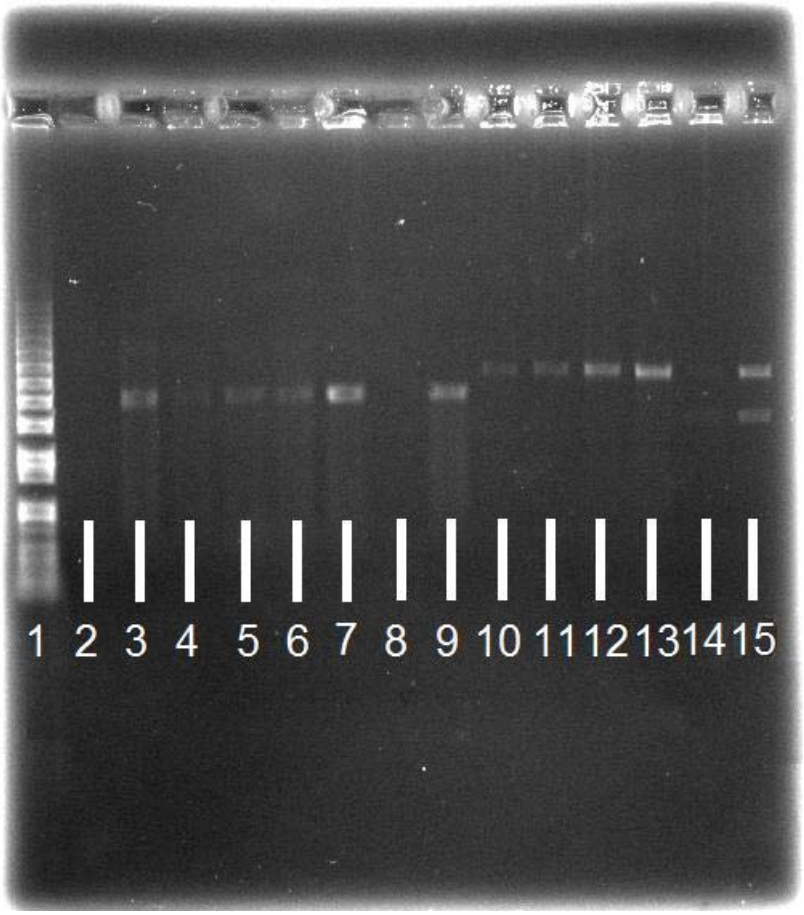


Figure 7.3: Gel Electrophoresis Image of the Longest Incubation Time for Each Temperature and Pressure. Lane 1, 1 kb+ ladder. Lanes 2-3, SS DNA at 0.1 MPa and 150 MPa respectively at 283.2 K. Lanes 4-7, SS DNA at 0.1 MPa, 50 MPa, 100 MPa and 150 MPa respectively at 313.2 K. Lanes 8-9, SS DNA at 0.1 MPa and 150 MPa respectively at 333.2 K. Lanes 10-13, RFII DNA at 0.1 MPa, 50 MPa, 100 MPa and 150 MPa respectively at 333.2 K. Lanes 14-15, RFI DNA at 0.1 MPa and 150 MPa respectively at 333.2 K.

From these results it was determined that optimal conditions for incubation times, considering time and the desired level of hydrolysis, were: 283.2 K at 2 weeks, 313.2 K at 1 week and 333.2 K for 1 day. These temperatures and times will provide a sufficient level of hydrolysis for the three temperature levels.

7.4.2 Post Amplification Gel Electrophoresis

After the amplification process, gel electrophoresis was used to assess the quality of each sample. The initial gel (Figure 7.4) resulted in clear bands for 48 of all 59 samples amplified. However, samples 5-9, 18-22 and 30 produced no bands (note that sample numbers 4 and 41 are missing. Miss-labelling of tubes resulted in the loss of these samples). Samples 5-9 correspond to the last 5 time points of the 283.2 K, 0.1 MPa incubation, samples 19-22 are all the samples from the 313.2 K, 0.1 MPa incubation and samples 19 and 30 are the last time points from the 283.2 K, 150 MPa and 313.2 K, 100 MPa incubations respectively.

This lack of bands has been attributed to an insufficient quantity of DNA present during amplification. To ensure the presence of DNA, an additional gel was run on the original samples of those which failed to amplify (Figure 7.5). Lanes 1-5 still show now signal while lanes 6-11 have bands. The samples corresponding to lanes 6-11 (samples 10, 19-22 and 30 from Figure 7.4 respectively) were amplified a second time. The lack of bands for samples corresponding to lanes 1- 5 (samples 5-9 from Figure 7.4 respectively) show that there is insufficient DNA for amplification. This has most likely resulted from DNA precipitation. These samples have been discarded.

The concentration of each DNA sample was determined with the use of a nanodrop spectrophotometer. The concentration of each sample is within the range of 6-60 ng μL^{-1}. This result is expected as the quantity of the amplified DNA is strongly dependent on the initial amount.

All gels have shown bands at approximately 4.9 kb which is the length of the expected PCR product.

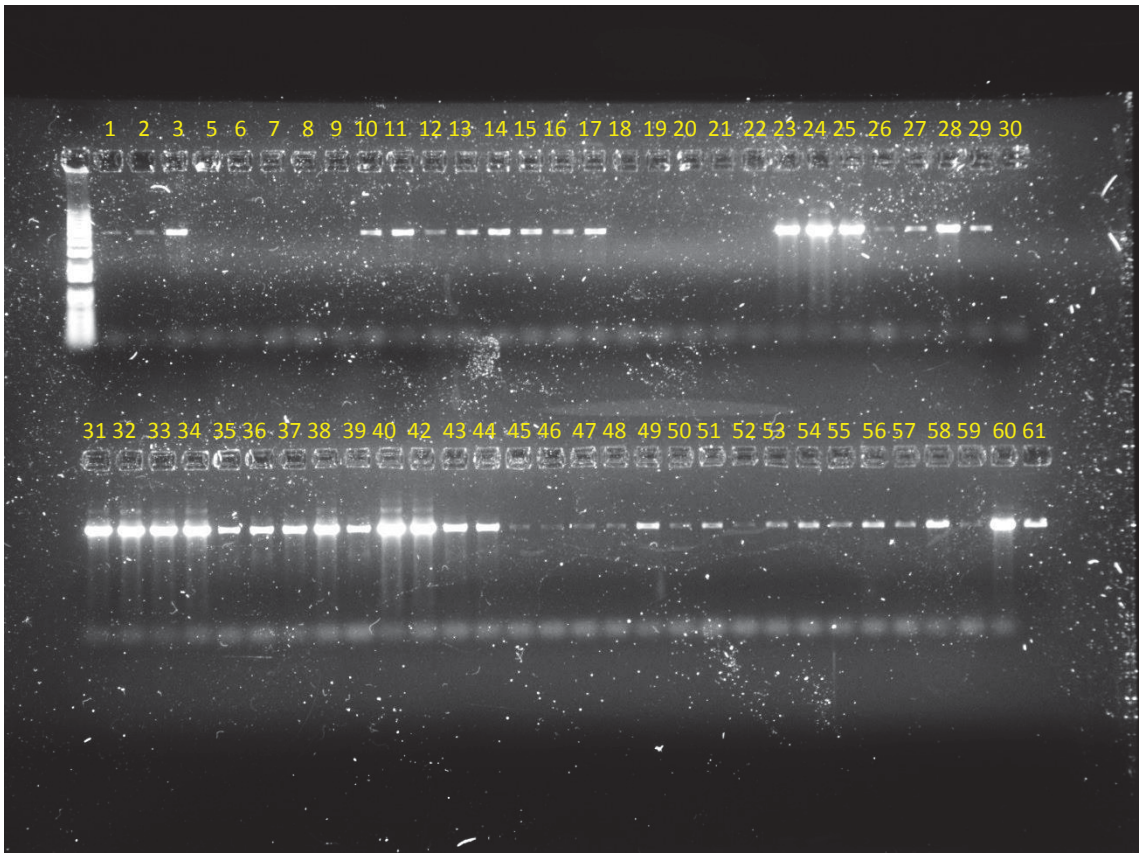


Figure 7.4: Post-Amplification Gel Electrophoresis Image. Lanes 1-9, 0.1 MPa, 283.2 K, SS DNA. Lanes 10-18, 150 MPa, 283.2 K, SS DNA. Lanes 19-22, 0.1 MPa, 313.2 K, SS DNA. Lanes 23-26, 50 MPa, 313.2 K, SS DNA. Lanes 19-22, 0.1 MPa, 313.2 K, SS DNA.

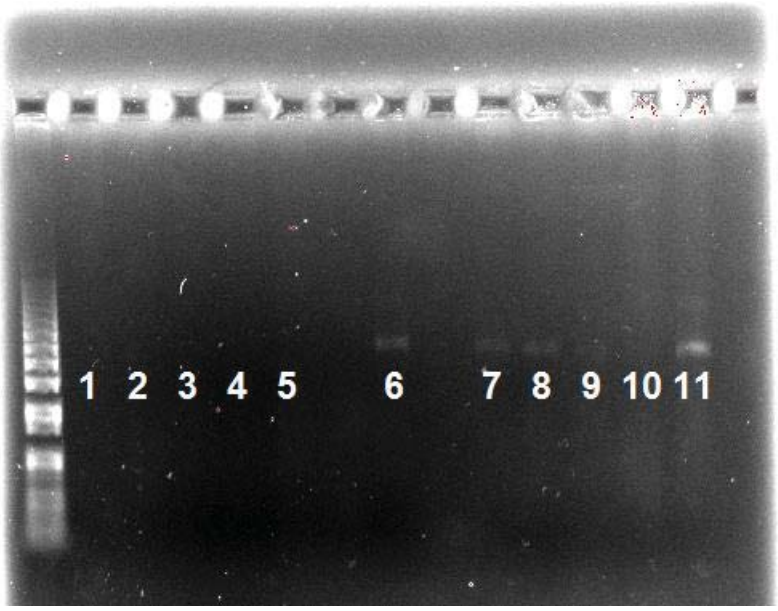


Figure 7.5: Gel Electrophoresis Image for Samples Which Failed to Amplify. Lanes correspond to samples 5-9,10, 19-22 and 30 from Figure 7.4 respectively.

7.4.3 DNA Sequencing

The time frame for the completion of sequencing provided by NZGL exceeds the completion date of this PhD and as a result, data on the hydrolysis of cytosine bases within ΦX174 bacteriophage have yet to be obtained. The estimated completion date of the sequencing by NZGL is 01/06/2015.

7.5 Discussion

In the absence of any sequencing results, a general outline of the information which is expected to be obtained for each incubation temperature is presented below.

7.5.1 283.2 K Incubation

Only two different incubation pressures were used at 283.2 K, 0.1 and 150 MPa. At this temperature and at 0.1 MPa, the rate constant for hydrolysis of cytosine within SS DNA is 6.81×10^{-13} s^{-1}. At this rate, the number of C-T transitions in 50,000 DNA copies for each individual cytosine much less than one over the two-week period. Because of this, it is impossible to gain data on the average behaviour of individual cytosines in this time frame. However, it is still possible to obtain data for the behaviour of the cytosines by studying the total number of cytosine C-T transitions observed in each chain for SS DNA. This can be used to compare to the average number of transitions observed in incubations of SS DNA at higher temperatures to help establish a rate of hydrolysis versus temperature curve for comparison to the results obtained by Frederico *et al*.

7.5.2 313.2 K Incubation

At 313.2 K, SS DNA was incubated at pressures of 0.1, 50, 100 and 150 MPa. At this temperature and at 0.1 MPa, the rate constant for hydrolysis of cytosine within SS DNA is 2.50 x 10^{-10} s^{-1}. At this rate, the number of C-T transitions over a one-week period should be sufficient to allow for data to be obtained for the rate of hydrolysis for each individual cytosine within the DNA sequence. With the number of pressure conditions used, it will be possible to examine the effects of high pressure on the rate of hydrolysis on each individual cytosine as a function of its environment.

7.5.3 333.2 K Incubation

At 333.2 K, RFII DNA was incubated at pressures of 0.1, 50, 100 and 150 MPa while single-stranded and RFI DNA were incubated at pressures of 0.1 and 150 MPa. At this temperature and at 0.1 MPa the rate constant for hydrolysis of cytosine within single-stranded DNA is 3.82 x 10^{-9} s^{-1} and for within double-stranded DNA is 2.73 x 10^{-11} s^{-1}. For the single-stranded DNA the number of C-T transitions over the one-day period will be more than sufficient to allow for data to be obtained for the rate of hydrolysis for each individual cytosine within the DNA sequence. These results combined with the other data from the lower temperatures will enable the construction of a pressure/temperature landscape for the average rate of hydrolysis of cytosine within single-stranded DNA. For the RFI and the RFII there will be insufficient data points to allow for the determination of the rate of hydrolysis for each individual cytosine. This is a result of the short incubation time required to prevent DNA degradation. Information on the average rate of hydrolysis will still be able to be obtained. A rate of hydrolysis versus pressure curve will be able to be established for the RFII DNA and this can be contrasted to the single-stranded DNA. There will only be two results for the RFI DNA but this will be sufficient to determine whether super-coiled DNA shows a different rate of hydrolysis compared to a nicked double-stranded sequence.

7.6 Conclusions

As it has been mentioned earlier, the chemical stability of cytosine is a limiting factor which must be taken into account when considering the conditions for the origin of life. A total of 58 DNA samples have been incubated at temperatures from 283.2 K to 333.2 K and at pressures of 0.1 MPa to 150MPa. These samples have then been amplified with a uracil-tolerant polymerase and have been submitted to NZ Genomics Ltd. for sequencing. Gel electrophoresis has shown that all samples submitted are clean with a length of approximately 4.9 kb. Once sequencing is complete, cytosine hydrolysis will be evident as a C → T transition. The sequencing results reveal the effect of high pressures on the chemical stability of cytosine as a function of pressure in single-stranded DNA, double-stranded nicked DNA and double-stranded super-coiled DNA. In addition to this, the sequencing results will also give insight as to whether the position of a cytosine residue in a DNA sequence will alter the effect of pressure on its rate of hydrolysis.

7.7 Future Work

Because of time constraints, this line of study has yet to be finished. Once the sequencing process is complete, data processing can begin. Once finished, the sequencing results will reveal how pressure affects the rate of hydrolysis of cytosine within a DNA sequence. In addition to finishing this work, there are a number of additional experiments which can be performed to expand upon this newly obtained knowledge. These experiments will be centred on effects of different chemical environments such as molecular crowding and the effects of salts.

7.7.1 Data Processing

Once sequencing is complete the resulting data will need to be processed. This will be a huge task with 64 different samples and 50,000 fold coverage for each. The sequencing data will be compared to the original DNA sequence allowing for the identification of any hydrolysed cytosines (visible as C to T transitions). The number and position of these transitions will be determined allowing for the identification of the average hydrolysis rate for each sample and in some cases the determination of the rate of hydrolysis for each individual base.

7.7.2 Other Conditions

Similar to the studies in **Chapter 3**, the studies performed so far have only involved solutions with a minimal presence of other solutes. However, this does not accurately reflect the conditions found outside of the lab. There have been considerations as to the effects of molecular crowding and other solutes. It is believed that molecular crowding may alter the rate of hydrolysis, while other solutes, such as amino acids, may have both stabilising and destabilising effects. This would be an in-depth study with multiple conditions to assess (concentrations of solutes, different molecular crowding agents, salts and amino acids) leading to a lengthy, expensive but worthwhile study.

Chapter 8 - Conclusions

High-pressure NMR techniques have been employed in order to further the current understanding of the effects of high pressure on the chemical and physical stability of DNA and its components. High-pressure studies were performed using a specially developed pressure system which provided high-pressure NMR capabilities along with a large volume high-pressure incubation vessel.

The rate of hydrolysis of cytosine and cytidine under high-pressure conditions has been studied with the use of high-pressure ^{1}H NMR spectroscopy. It has been observed that the rates of hydrolysis of both cytosine and cytidine at 373.2 K increase significantly at pressures of 150 MPa compared to those at atmospheric pressure (0.42×10^{-6} s^{-1} to 0.72×10^{-6} s^{-1} and 0.44×10^{-6} s^{-1} to 0.87×10^{-6} s^{-1} respectively). It was also possible to determine the reaction volumes of these nucleobases (-11.7 ± 1.2 cm^3 mol^{-1} and -14.6 ± 1.9 cm^3 mol^{-1} respectively). The rate of hydrolysis of cytosine and cytidine, which was already considered to be rather fast on the geological time scale, has been further increased by pressure, thus reducing the chemical stability of these nucleobases. This result disfavours high-temperature/high-pressure origin-of-life theories which involve cytosine in the DNA/RNA structure.

High-pressure ^{1}H NMR spectroscopy was also used to study the thermal stability of hexameric, dodecameric and i-motif DNA sequences. It was observed that the thermal stability of a small, self-complementary, DNA hexamer decreased marginally under high pressure conditions (a decrease of 7.4 K from 0.1 MPa to 240 MPa). This result is believed to be largely influenced by the size of the DNA.

Further studies using self-complementary DNA dodecamers have shown that the effect of high pressures on the thermal stability of DNA is rather complex. DNA sequences with a larger A-T content were observed to have slightly higher melting points under high-pressure conditions (an increase of approximately 1.6 K from 0.1 MPa to 200 MPa), whereas sequences with high G-C content were observed to first have a decreased melting point at 50 MPa compared to 0.1 MPa followed by

an increased melting point at pressures above 100 MPa. These results show a significant correlation to results obtained by Hedwig *et al.* [81-83] on the partial molar isothermal compressions of the four major nucleobases.

The effects of pressure on the thermal stability of an i-motif structure were also studied under different pH conditions with the use of high-pressure ^{1}H NMR spectroscopy. These results show that an i-motif has less thermal stability under high pressures but only when pressure-induced pH shifts are properly taken into account. This change in thermal stability is a result of the pK_a of cytosine shifting under high-pressure conditions.

These results on the thermal stability of DNA sequences are varied, with both positive and negative pressure-induced effects. However, it can be concluded that, at pressures in the range of 50 MPa to 100 MPa, the thermal stabilities of hexamers, dodecamers with high G-C content and i-motif structures are lower than at atmospheric pressure.

Samples of ΦX174 bacteriophage were incubated under high pressure conditions using the high-pressure reaction vessel. These samples were then amplified using a uracil tolerant polymerase and submitted to NZ Genomics for sequencing using Illumina SBS technology. The sequencing of these samples has yet to be completed, but will be used to examine the effects of high pressure on the chemical stability of cytosine within both single-stranded and double-stranded DNA structures.

The results obtained over the course of this study have provided insight in to the effects of high pressures on the stability, both chemical and physical, of DNA and its components. While these results cover various possibilities, it is clear that a high-pressure origin-of-life involving cytosine is strongly disfavoured on many grounds. The rate of hydrolysis of cytosine, by far the least stable of the major nucleobases, is increased under high-pressure conditions, while the thermal stability of DNA dodecamers with a high G-C content is lower at pressures up to 100 MPa (oceanic pressures). These combined decreases in both chemical and physical stabilities under high-pressure conditions decrease the likelihood of a high-pressure/high-temperature origin-of-life involving cytosine.

Bibliography

1. Orgel, L.E., *Prebiotic chemistry and the origin of the RNA world.* Crit Rev Biochem Mol Biol, 2004. **39**(2): p. 99-123.

2. Bartel, D.P. and P.J. Unrau, *Constructing an RNA world.* Trends in Biochemical Sciences, 1999. **24**(12): p. M9-M13.

3. Andras, P. and C. Andras, *The origins of life – the 'protein interaction world' hypothesis: protein interactions were the first form of self-reproducing life and nucleic acids evolved later as memory molecules.* Medical Hypotheses, 2005. **64**(4): p. 678-688.

4. Nelson, K.E., M. Levy, and S.L. Miller, *Peptide nucleic acids rather than RNA may have been the first genetic molecule.* Proceedings of the National Academy of Sciences of the United States of America, 2000. **97**(8): p. 3868-3871.

5. Forterre, P., N. Benachenhou-Lahfa, F. Confalonieri, M. Duguet, C. Elie, and B. Labedan, *The nature of the last universal ancestor and the root of the tree of life, still open questions.* Biosystems, 1992. **28**(1-3): p. 15-32.

6. Woese, C.R., O. Kandler, and M.L. Wheelis, *Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya.* Proceedings of the National Academy of Sciences, 1990. **87**(12): p. 4576-4579.

7. Penny, D. and A. Poole, *The nature of the last universal common ancestor.* Curr Opin Genet Dev, 1999. **9**(6): p. 672-7.

8. Doherty, E.A. and J.A. Doudna, *Ribozyme structures and mechanisms.* Annu Rev Biophys Biomol Struct, 2001. **30**: p. 457-75.

9. Dworkin, J.P., A. Lazcano, and S.L. Miller, *The roads to and from the RNA world.* J Theor Biol, 2003. **222**(1): p. 127-34.

10. Vicens, Q. and T.R. Cech, *A natural ribozyme with 3[prime],5[prime] RNA ligase activity.* Nat Chem Biol, 2009. **5**(2): p. 97-99.

11. Hayden, E.J., G. von Kiedrowski, and N. Lehman, *Systems chemistry on ribozyme self-construction: evidence for anabolic autocatalysis in a recombination network.* Angew Chem Int Ed Engl, 2008. **47**(44): p. 8424-8.

12. Voytek, S.B. and G.F. Joyce, *Niche partitioning in the coevolution of 2 distinct RNA enzymes.* Proceedings of the National Academy of Sciences, 2009. **106**(19): p. 7780-7785.

13. Ashkenazy, Y., H. Gildor, M. Losch, F.A. Macdonald, D.P. Schrag, and E. Tziperman, *Dynamics of a Snowball Earth ocean.* Nature, 2013. **495**(7439): p. 90-93.

14. Sinha, R. and K. Krishnan, *Novel opportunity for understanding origin and evolution of life: perspectives on the exploration of subglacial environment of Lake Vostok, Antarctica.* Annals of Microbiology, 2013. **63**(2): p. 409-415.

15. Vlassov, A.V., B.H. Johnston, L.F. Landweber, and S.A. Kazakov, *RNA Catalysis in Frozen Solutions.* Doklady Biochemistry & Biophysics, 2005. **402**(1-6): p. 207-209.

16. Vlassov, A.V., B.H. Johnston, L.F. Landweber, and S.A. Kazakov, *Ligation activity of fragmented ribozymes in frozen solution: implications for the RNA world.* Nucleic acids research, 2004. **32**(9): p. 2966-2974.

17. Kazakov, S.A., S.V. Balatskaya, and B.H. Johnston, *Ligation of the hairpin ribozyme in cis induced by freezing and dehydration.* Rna, 2006. **12**(3): p. 446-56.

18. Bada, J.L. and A. Lazcano, *Origin of life. Some like it hot, but not the first biomolecules.* Science, 2002. **296**(5575): p. 1982-3.

19. Koonin, E.V. and W. Martin, *On the origin of genomes and cells within inorganic compartments.* Trends Genet, 2005. **21**(12): p. 647-54.

20. Obscura, A., *http://www.atlasobscura.com/places/fly-ranch-geyser.* 2014.

21. Planet, A., *http://www.amusingplanet.com/2013/05/the-majestic-tufa-towers-of-mono-lake.html.* 2015.

22. Segerer, A.H., S. Burggraf, G. Fiala, G. Huber, R. Huber, U. Pley, and K.O. Stetter, *Life in hot springs and hydrothermal vents.* Orig Life Evol Biosph, 1993. **23**(1): p. 77-90.

23. Tarasov, V.G., A.V. Gebruk, A.N. Mironov, and L.I. Moskalev, *Deep-sea and shallow-water hydrothermal vent communities: Two different phenomena?* Chemical Geology, 2005. **224**(1–3): p. 5-39.

24. Haase, K.M., S. Petersen, A. Koschinsky, R. Seifert, C.W. Devey, R. Keir, K.S. Lackschewitz, B. Melchert, M. Perner, O. Schmale, J. Süling, N. Dubilier, F. Zielinski, S. Fretzdorff, D. Garbe-Schönberg, U. Westernströer, C.R. German, T.M. Shank, D. Yoerger, O. Giere, J. Kuever, H. Marbler, J. Mawick, C. Mertens, U. Stöber, M. Walter, C. Ostertag-Henning, H. Paulick, M. Peters, H. Strauss, S. Sander, J. Stecher, M. Warmuth, and S. Weber, *Young volcanism and related hydrothermal activity at 5°S on the slow-spreading southern Mid-Atlantic Ridge.* Geochemistry, Geophysics, Geosystems, 2007. **8**(11): p. Q11002.

25. Vieille, C. and G.J. Zeikus, *Hyperthermophilic Enzymes: Sources, Uses, and Molecular Mechanisms for Thermostability.* Microbiology and Molecular Biology Reviews, 2001. **65**(1): p. 1-43.

26. Vergne, J., J.A.H. Cognet, E. Szathmáry, and M.-C. Maurel, *In vitro selection of halo-thermophilic RNA reveals two families of resistant RNA.* Gene, 2006. **371**(2): p. 182-193.

27. Levy, M. and S.L. Miller, *The stability of the RNA bases: Implications for the origin of life.* Proceedings of the National Academy of Sciences, 1998. **95**(14): p. 7933-7938.

28. Frederico, L.A., T.A. Kunkel, and B.R. Shaw, *A sensitive genetic assay for the detection of cytosine deamination: determination of rate constants and the activation energy.* Biochemistry, 1990. **29**(10): p. 2532-7.

29. Kawamura, K., *Behaviour of RNA under hydrothermal conditions and the origins of life.* International Journal of Astrobiology, 2004. **3**(04): p. 301-309.

30. Moulton, V., P.P. Gardner, R.F. Pointon, L.K. Creamer, G.B. Jameson, and D. Penny, *RNA Folding Argues Against a Hot-Start Origin of Life.* Journal of Molecular Evolution, 2000. **51**(4): p. 416-421.

31. Fedoruk-Wyszomirska, A., E. Wyszko, M. Giel-Pietraszuk, M.Z. Barciszewska, and J. Barciszewski, *High hydrostatic pressure approach proves RNA catalytic activity without magnesium.* Int J Biol Macromol, 2007. **41**(1): p. 30-5.

32. Krzyżaniak, A., P. Sałański, R.W. Adamiak, J. Jurczak, and J. Barciszewski, *Structure and Function of Nucleic Acids Under High Pressure*, in *Progress in Biotechnology*, R. Hayashi and C. Balny, Editors. 1996, Elsevier. p. 189-194.

33. Biao, L., L. Qing, L. Wangyi, and R. Kangcheng, *Stability of the Sarcin/Ricin domain of 28S RNA in rat ribosome under hydrostatic pressure.* Chinese Science Bulletin, 1998. **43**(4): p. 303-305.

34. Wilton, D.J., M. Ghosh, K.V. Chary, K. Akasaka, and M.P. Williamson, *Structural change in a B-DNA helix with hydrostatic pressure.* Nucleic Acids Res, 2008. **36**(12): p. 4032-7.

35. Mombelli, E., E. Shehi, P. Fusi, and P. Tortora, *Exploring hyperthermophilic proteins under pressure: theoretical aspects and experimental findings.* Biochimica et Biophysica Acta (BBA) - Protein Structure and Molecular Enzymology, 2002. **1595**(1–2): p. 392-396.

36. Marqués, M.I., J.M. Borreguero, H.E. Stanley, and N.V. Dokholyan, *Possible Mechanism for Cold Denaturation of Proteins at High Pressure.* Physical Review Letters, 2003. **91**(13): p. 138103.

37. Yayanos, A., *Advances in High Pressure Bioscience and Biotechnology*, ed. H. Ludwig. 1999, Berlin: Springer. 3.

38. Nash, D.P. and J. Jonas, *Structure of Pressure-Assisted Cold Denatured Lysozyme and Comparison with Lysozyme Folding Intermediates.* Biochemistry, 1997. **36**(47): p. 14375-14383.

39. Li, H., H. Yamada, and K. Akasaka, *Effect of pressure on individual hydrogen bonds in proteins. Basic pancreatic trypsin inhibitor.* Biochemistry, 1998. **37**(5): p. 1167-73.

40. Kuwata, K., H. Li, H. Yamada, C.A. Batt, Y. Goto, and K. Akasaka, *High pressure NMR reveals a variety of fluctuating conformers in β-lactoglobulin.* Journal of Molecular Biology, 2001. **305**(5): p. 1073-1083.

41. Conradi, M.S., *High-Pressure NMR: Techniques*, in *Encyclopedia of Magnetic Resonance*. 2007, John Wiley & Sons, Ltd.

42. Yonker, C.R., T.S. Zemanian, S.L. Wallen, J.C. Linehan, and J.A. Franz, *A New Apparatus for the Convenient Measurement of NMR Spectra in High-Pressure Liquids.* Journal of Magnetic Resonance, Series A, 1995. **113**(1): p. 102-107.

43. Yamada, H., M. Nakatsuka, H. Yamochi, M. Sawamura, and A. Sera, *Spinning glass cell for high-pressure high-resolution NMR measurements.* Review of Scientific Instruments, 1991. **62**(3): p. 700-704.

44. Jonas, J., L. Ballard, and D. Nash, *High-resolution, high-pressure NMR studies of proteins.* Biophys J, 1998. **75**(1): p. 445-52.

45. Daedalus Innovations. *http://www.daedalusinnovations.com/*. 2007.

46. Christopher Roe, D., *Sapphire NMR tube for high-resolution studies at elevated pressure.* Journal of Magnetic Resonance (1969), 1985. **63**(2): p. 388-391.

47. Erlach, M.B., C.E. Munte, W. Kremer, R. Hartl, D. Rochelt, D. Niesner, and H.R. Kalbitzer, *Ceramic cells for high pressure NMR spectroscopy of proteins.* Journal of Magnetic Resonance, 2010. **204**(2): p. 196-199.

48. HiP. *http://www.highpressure.com/*. 2015.

49. Bruins, M.E., A.M. Matser, A.E.M. Janssen, and R.M. Boom, *Buffer selection for HP treatment of biomaterials and its consequences for enzyme inactivation studies.* International Journal of High Pressure Research, 2007. **27**(1): p. 101-107.

50. El'yanov, B. and S. Hamann, *Some quantitative relationships for ionization reactions at high pressures.* Australian Journal of Chemistry, 1975. **28**(5): p. 945-954.

51. Kitamura, Y. and T. Itoh, *Reaction volume of protonic ionization for buffering agents. Prediction of pressure dependence of pH and pOH.* Journal of Solution Chemistry, 1987. **16**(9): p. 715-725.

52. Lazcano, A. and S. Miller, *How long did it take for life to begin and evolve to cyanobacteria?* Journal of Molecular Evolution, 1994. **39**(6): p. 546-554.

53. Piccirilli, J.A., S.A. Benner, T. Krauch, and S.E. Moroney, *Enzymatic incorporation of a new base pair into DNA and RNA extends the genetic alphabet.* Nature, 1990. **343**(6253): p. 33-37.

54. Crick, F.H.C., *The origin of the genetic code.* Journal of Molecular Biology, 1968. **38**(3): p. 367-379.

55. Hazen, R.M., N. Boctor, J.A. Brandes, G.D. Cody, R.J. Hemley, A. Sharma, and H.S.Y. Jr, *High pressure and the origin of life.* Journal of Physics: Condensed Matter, 2002. **14**(44): p. 11489.

56. Almatarneh, M.H., C.G. Flinn, and R.A. Poirier, *Mechanisms for the Deamination Reaction of Cytosine with H2O/OH− and 2H2O/OH−: A Computational Study.* Journal of Chemical Information and Modeling, 2008. **48**(4): p. 831-843.

57. Almatarneh, M.H., C.G. Flinn, R.A. Poirier, and W.A. Sokalski, *Computational Study of the Deamination Reaction of Cytosine with H2O and OH.* The Journal of Physical Chemistry A, 2006. **110**(26): p. 8227-8234.

58. Notari, R.E., *A mechanism for the hydrolytic deamination of cytosine arabinoside in aqueous buffer.* J Pharm Sci, 1967. **56**(7): p. 804-9.

59. Garrett, E.R. and J. Tsau, *Solvolyses of cytosine and cytidine.* J Pharm Sci, 1972. **61**(7): p. 1052-61.

60. Mozhaev, V.V., K. Heremans, J. Frank, P. Masson, and C. Balny, *High pressure effects on protein structure and function.* Proteins: Structure, Function, and Bioinformatics, 1996. **24**(1): p. 81-91.

61. *The World Ocean*, in *The Columbia Encyclopedia*. 2007, Columbia University Press.

62. *IHO-IOC GEBCO Gazetteer of Undersea Feature Names, August 2011 version*. 2011, GEBCO.

63. White, R.H., *Hydrolytic stability of biomolecules at high temperatures and its implication for life at 250 degrees C.* Nature, 1984. **310**(5976): p. 430-2.

64. Wojciechowski, F. and C.J. Leumann, *Alternative DNA base-pairs: from efforts to expand the genetic code to potential material applications.* Chem Soc Rev, 2011. **40**(12): p. 5669-79.

65. Moulton, V., P.P. Gardner, R.F. Pointon, L.K. Creamer, G.B. Jameson, and D. Penny, *RNA folding argues against a hot-start origin of life.* J Mol Evol, 2000. **51**(4): p. 416-21.

66. Refaee, M., T. Tezuka, K. Akasaka, and M.P. Williamson, *Pressure-dependent changes in the solution structure of hen egg-white lysozyme.* J. Mol. Biol., 2003. **327**(4): p. 857-865.

67. Jonas, J. and A. Jonas, *High-Pressure NMR Spectroscopy of Proteins and Membranes.* Annual Review of Biophysics and Biomolecular Structure, 1994. **23**(1): p. 287-318.

68. Wilton, D.J., R.B. Tunnicliffe, Y.O. Kamatari, K. Akasaka, and M.P. Williamson, *Pressure-induced changes in the solution structure of the GB1 domain of protein G.* Proteins: Structure, Function, and Bioinformatics, 2008. **71**(3): p. 1432-1440.

69. Gross, M. and R. Jaenicke, *Proteins under pressure.* European Journal of Biochemistry, 1994. **221**(2): p. 617-630.

70. Williamson, M.P., K. Akasaka, and M. Refaee, *The solution structure of bovine pancreatic trypsin inhibitor at high pressure.* Protein Sci, 2003. **12**(9): p. 1971-9.

71. Krzyźaniak, A., P. Salański, J. Jurczak, and J. Barciszewski, *B • Z DNA reversible conformation changes effected by high pressure.* FEBS Letters, 1991. **279**(1): p. 1-4.

72. Wu, J.Q. and R.B. Macgregor, Jr., *Pressure dependence of the helix-coil transition temperature of poly[d(G-C)].* Biopolymers, 1995. **35**(4): p. 369-76.

73. Wu, J.Q. and R.B. Macgregor, Jr., *Pressure dependence of the melting temperature of dA.dT polymers.* Biochemistry, 1993. **32**(46): p. 12531-7.

74. Owczarzy, R., B.G. Moreira, Y. You, M.A. Behlke, and J.A. Walder, *Predicting Stability of DNA Duplexes in Solutions Containing Magnesium and Monovalent Cations.* Biochemistry, 2008. **47**(19): p. 5336-5353.

75. Doktycz, M.J., *Nucleic Acids: Thermal Stability and Denaturation*, in *eLS*. 2001, John Wiley & Sons, Ltd.

76. Hwang, T.L. and A.J. Shaka, *Water Suppression That Works. Excitation Sculpting Using Arbitrary Wave-Forms and Pulsed-Field Gradients.* Journal of Magnetic Resonance, Series A, 1995. **112**(2): p. 275-279.

77. Nielsen, C.B., M. Petersen, E.B. Pedersen, P.E. Hansen, and U.B. Christensen, *NMR Structure Determination of a Modified DNA Oligonucleotide Containing a New Intercalating Nucleic Acid.* Bioconjugate Chemistry, 2004. **15**(2): p. 260-269.

78. Wuthrich, K., *NMR of Proteins and NUcleic Acids*. 1986: Wiley-Interscience.

79. Gnuplot. *http://www.gnuplot.info/*. 2015.

80. Kool, E.T., *Hydrogen Bonding, Base Stacking, and Steric Effects in DNA Replication.* Annual Review of Biophysics & Biomolecular Structure, 2001. **30**(1): p. 1-22.

81. Hedwig, G., G. Jameson, and H. Høiland, *Volumetric Properties of the Nucleoside Thymidine in Aqueous Solution at T = 298.15 K and p = (10 to 100) MPa.* Journal of Solution Chemistry, 2014. **43**(4): p. 804-820.

82. Hedwig, G.R., E. Høgseth, and H. Høiland, *Volumetric properties of the nucleosides adenosine, cytidine, and uridine in aqueous solution at T = 298.15 K and p = (10 to 120) MPa.* The Journal of Chemical Thermodynamics, 2013. **61**(0): p. 117-125.

83. Hedwig, G.R., G.B. Jameson, and H. Høiland, *Volumetric Properties at High Pressures of the Nucleosides Inosine, 2′-Deoxyinosine, and 2′-Deoxyguanosine and the Volumetric Properties of Guanosine Derived Using Group Additivity Methods.* Journal of Chemical & Engineering Data, 2014. **59**(11): p. 3593-3604.

84. Ghosh, A. and M. Bansal, *A glossary of DNA structures from A to Z.* Acta Crystallogr D Biol Crystallogr, 2003. **59**(Pt 4): p. 620-6.

85. Htun, H. and J. Dahlberg, *Single strands, triple strands, and kinks in H-DNA.* Science, 1988. **241**(4874): p. 1791-1796.

86. Lim, K.W. and A.T. Phan, *Structural Basis of DNA Quadruplex–Duplex Junction Formation.* Angewandte Chemie International Edition, 2013. **52**(33): p. 8566-8569.

87. Burge, S., G.N. Parkinson, P. Hazel, A.K. Todd, and S. Neidle, *Quadruplex DNA: sequence, topology and structure.* Nucleic acids research, 2006. **34**(19): p. 5402-5415.

88. Collie, G.W. and G.N. Parkinson, *The application of DNA and RNA G-quadruplexes to therapeutic medicines.* Chem Soc Rev, 2011. **40**(12): p. 5867-92.

89. Catasti, P., X. Chen, L.L. Deaven, R.K. Moyzis, E.M. Bradbury, and G. Gupta, *Cystosine-rich strands of the insulin minisatellite adopt hairpins with intercalated Cytosine+·Cytosine pairs.* Journal of Molecular Biology, 1997. **272**(3): p. 369-382.

90. Blackburn, E.H. and J.G. Gall, *A tandemly repeated sequence at the termini of the extrachromosomal ribosomal RNA genes in Tetrahymena.* J Mol Biol, 1978. **120**(1): p. 33-53.

91. Blackburn, E.H., *Structure and function of telomeres.* Nature, 1991. **350**(6319): p. 569-573.

92. Blackburn, E., *Telomeres and their synthesis.* Science, 1990. **249**(4968): p. 489-490.

93. Brown, W.R., P.J. MacKinnon, A. Villasante, N. Spurr, V.J. Buckle, and M.J. Dobson, *Structure and polymorphism of human telomere-associated DNA.* Cell, 1990. **63**(1): p. 119-32.

94. Murray, A.W. and J.W. Szostak, *Construction of artificial chromosomes in yeast.* Nature, 1983. **305**(5931): p. 189-193.

95. Chan, Y.-C., J.-W. Chen, S.-Y. Su, and C.-C. Chang, *Direct visualization of the quadruplex structures in human chromosome using FRET: Application of quadruplex stabilizer and duplex-binding fluorophore.* Biosensors and Bioelectronics, 2013. **47**(0): p. 566-573.

96. Školáková, P., K. Bednářová, M. Vorlíčková, and J. Sagi, *Quadruplexes of human telomere dG3(TTAG3)3 sequences containing guanine abasic sites.* Biochemical and Biophysical Research Communications, 2010. **399**(2): p. 203-208.

97. Kaushik, M., M. Prasad, S. Kaushik, A. Singh, and S. Kukreti, *Structural transition from dimeric to tetrameric i-motif, caused by the presence of TAA at the 3′-end of human telomeric C-rich sequence.* Biopolymers, 2010. **93**(2): p. 150-160.

98. Kaushik, M., N. Suehl, and L.A. Marky, *Calorimetric unfolding of the bimolecular and i-motif complexes of the human telomere complementary strand, d(C3TA2)4.* Biophysical Chemistry, 2007. **126**(1–3): p. 154-164.

99. Zhao, Y., Z.-y. Kan, Z.-x. Zeng, Y.-h. Hao, H. Chen, and Z. Tan, *Determining the Folding and Unfolding Rate Constants of Nucleic Acids by Biosensor. Application to Telomere G-Quadruplex.* Journal of the American Chemical Society, 2004. **126**(41): p. 13255-13264.

100. Leroy, J.-L., M. Guéron, J.-L. Mergny, and C. Hélène, *Intramolecular folding of a fragment of the cytosine-rich strand of telomeric DNA into an i-motif.* Nucleic acids research, 1994. **22**(9): p. 1600-1606.

101. Takahashi, S. and N. Sugimoto, *Effect of Pressure on the Stability of G-Quadruplex DNA: Thermodynamics under Crowding Conditions.* Angewandte Chemie International Edition, 2013. **52**(51): p. 13774-13778.

102. Fan, H.Y., Y.L. Shek, A. Amiri, D.N. Dubins, H. Heerklotz, R.B. Macgregor, and T.V. Chalikian, *Volumetric Characterization of Sodium-Induced G-Quadruplex Formation.* Journal of the American Chemical Society, 2011. **133**(12): p. 4518-4526.

103. Modi, S., A.H. Wani, and Y. Krishnan, *The PNA–DNA hybrid I-motif: implications for sugar–sugar contacts in i-motif tetramerization.* Nucleic acids research, 2006. **34**(16): p. 4354-4363.

104. Krishnan-Ghosh, Y., E. Stephens, and S. Balasubramanian, *PNA forms an i-motif.* Chem Commun, 2005. **14**(42): p. 5278-80.

105. Future, D. *http://www.drugfuture.com/chemdata/cytosine.html*. 2015.

106. Sanger, F., A.R. Coulson, T. Friedmann, G.M. Air, B.G. Barrell, N.L. Brown, J.C. Fiddes, C.A. Hutchison Iii, P.M. Slocombe, and M. Smith, *The nucleotide sequence of bacteriophage φX174.* Journal of Molecular Biology, 1978. **125**(2): p. 225-246.

107. Seo, T.S., X. Bai, D.H. Kim, Q. Meng, S. Shi, H. Ruparel, Z. Li, N.J. Turro, and J. Ju, *Four-color DNA sequencing by synthesis on a chip using photocleavable fluorescent nucleotides.* Proceedings of the National Academy of Sciences, 2005. **102**(17): p. 5926-5931.

108. Illumina, *http://www.illumina.com/documents/products/techspotlights/techspotlight_sequencing.pdf.* 2014.

Appendices

Appendix one – Cytosine/Cytidine Hydrolysis Data

Average cytosine/cytidine hydrolysis rate constants, *k*, for each pH and pressure condition.

Solution	*k* (0.1 MPa) /x 10^{-6} s^{-1}	*k* (50 MPa) /x 10^{-6} s^{-1}	*k* (100 MPa) /x 10^{-6} s^{-1}	*k* (150 MPa) /x 10^{-6} s^{-1}	*k* (200 MPa) /x 10^{-6} s^{-1}
Cytosine pH 6.0	0.592 ± 0.067			0.866 ± 0.069	
Cytosine pH 7.0	0.416 ± 0.039	0.488 ± 0.036	0.615 ± 0.035	0.721 ± 0.063	0.878 ±0.042
Cytosine pH 8.0	0.424 ± 0.025			0.776 ± 0. 039	
Cytidine pH 7.0	0.440 ± 0.042			0.867 ± 0.083	

Appendix Two – Hexamer Melting Data

Hexamer melting points for each base proton under high-pressure conditions. Melting points determined by fitting of an empirical function are indicated by a * and those which are rough estimates due to large gaps in the data are indicated by a ***. All other melting points have been determined by fitting with equation (23).

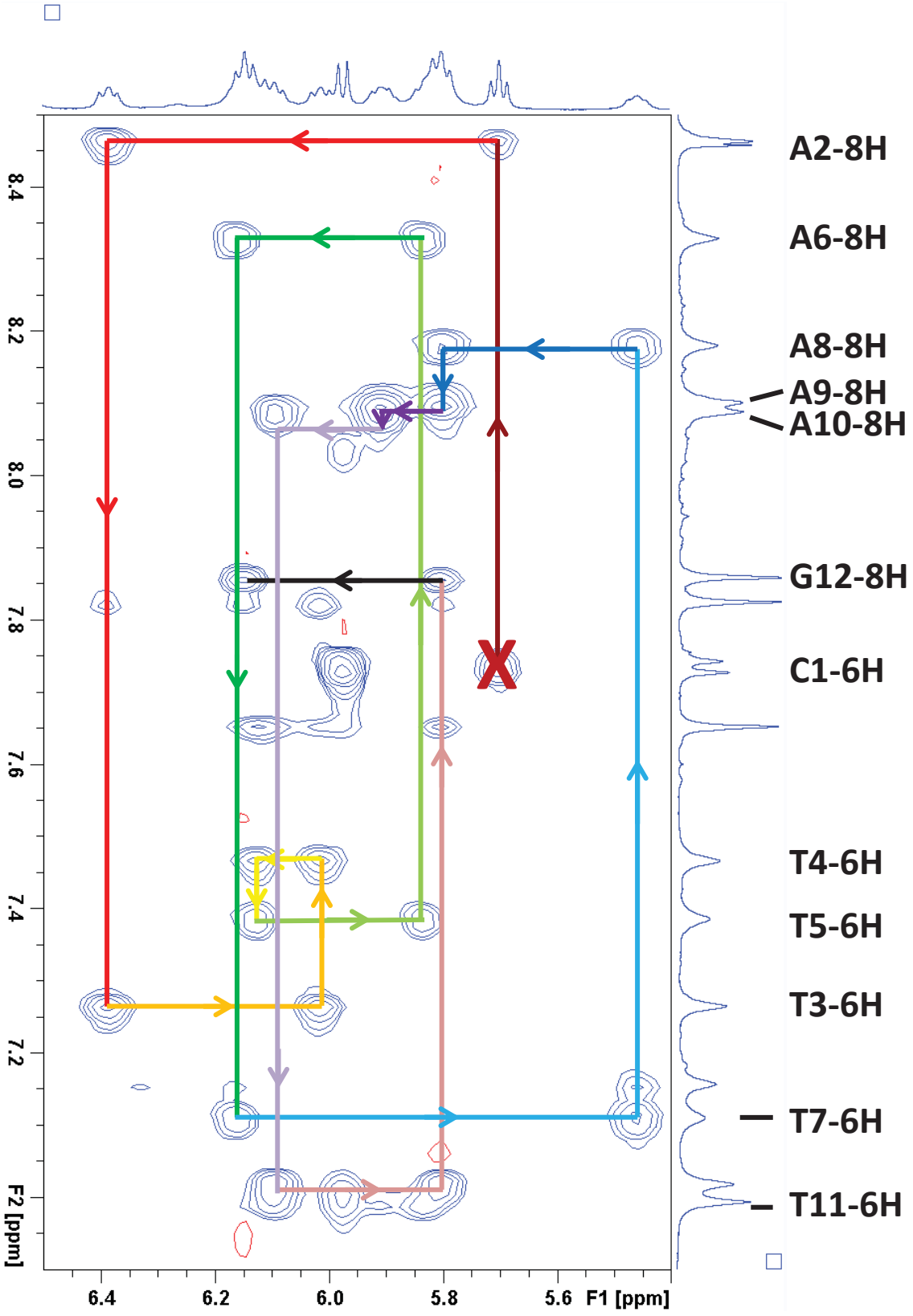
Proton	0.1 MPa	100 MPa	150 MPa	200 MPa	240 MPa
C1-6H*	292.72	290.74	289.55	288.39	287.00
A2-8H	298.79	294.39	292.78	290.65	289.24
T3-6H	296.75	296.31	296.48	294.12	293.42
A4-8H	302.06	301.00	296.22	292.25	291.94
T5-6H	293.97	292.17	291.49	290.68	289.04
G6-8H	297.03	297.04	296.31	294.86	292.00
A2-2H	294.23	292.77	289.57	289.35	287.37
A4-2H* **	295.3	290.0	287.5	285.3	280.2
C1-5H	295.18	293.28	292.20	292.16	290.28
T5-3CH_3*	295.14	292.90	292.06	289.61	288.80
T5-5CH_3	297.38	295.92	293.43	288.92	287.62
Average	296.23	294.22	292.51	290.57	288.81

Appendix Three – Dodecamer Melting Data

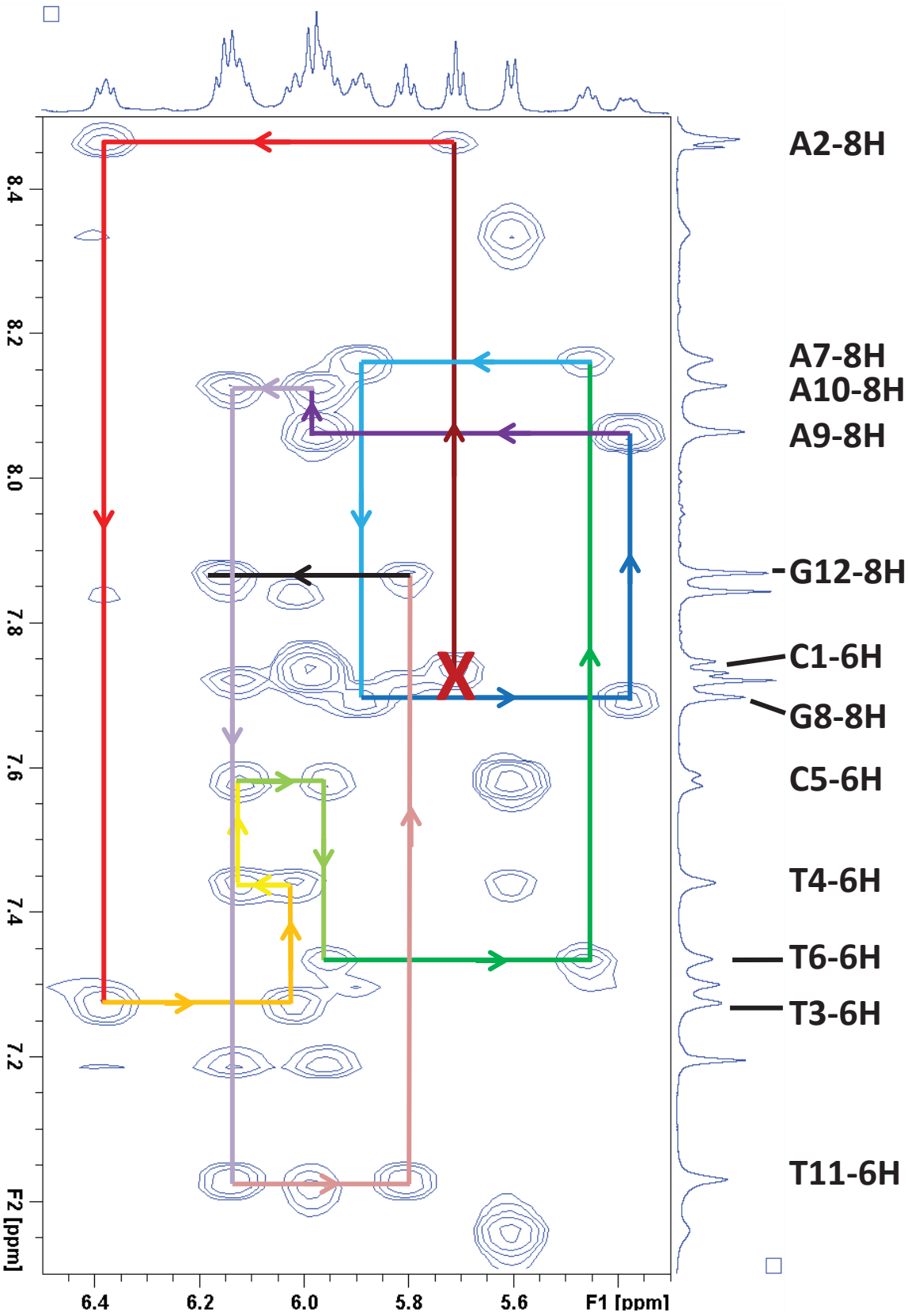
Spectral Assignment

Signal assignment link pathways for dodecamer DNA sequences 121 (5′-CAT TTA TAA ATG-3′), 122 (5′-CAT TCT AGA ATG-3′), 124 (5′-CAG GTC GAC CTG-3′) and 125 (5′-CAC CCG CGG GTG-3′), with d_i(6H/8H→1′H) and d_s(1′H→6H/8H) connectivities. The link pathway can be followed from the dark red X along the lines in "rainbow" order through to the final black line. Unlabelled peaks are attributed A-2H signals and were not used to establish links.

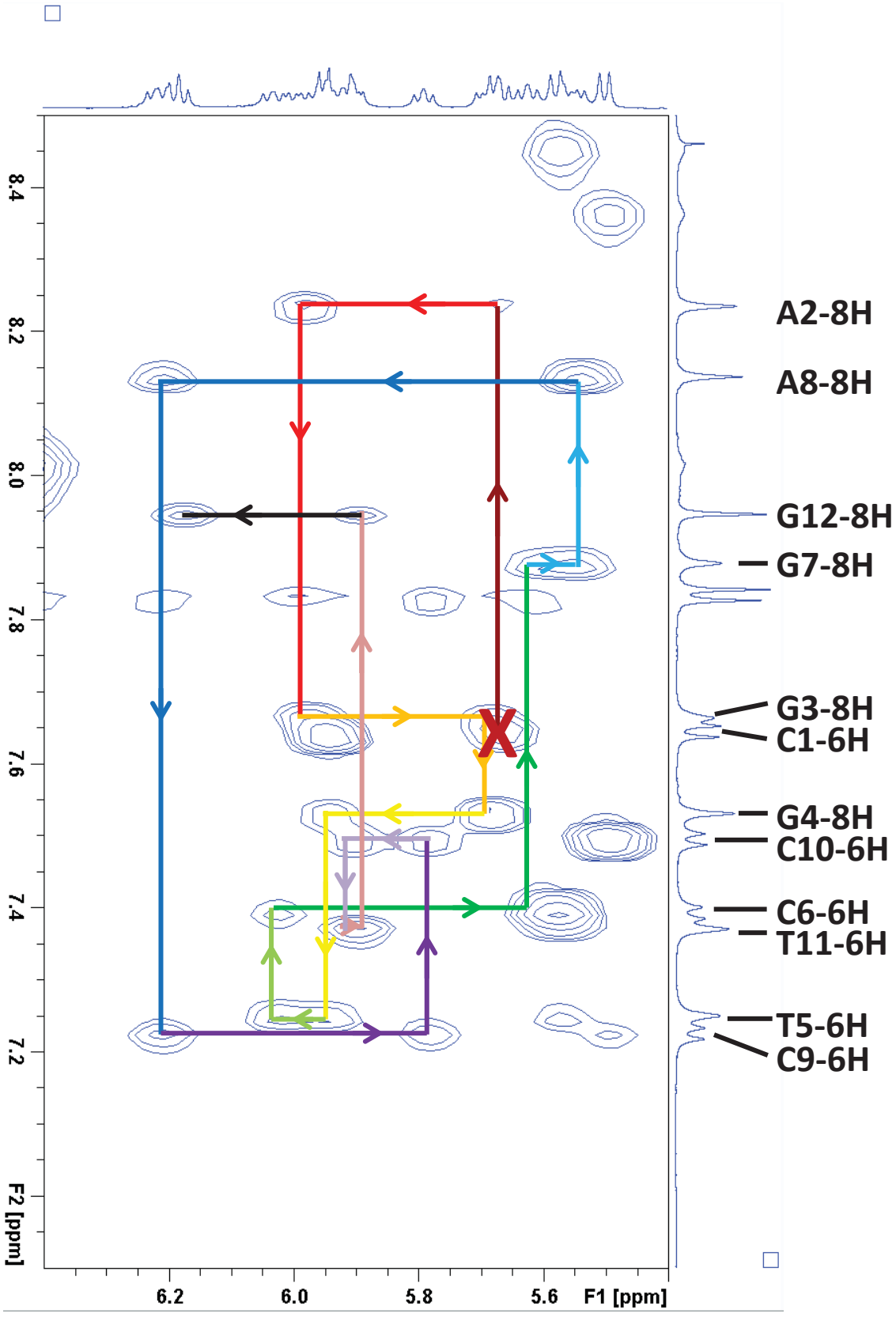
Sequence 121



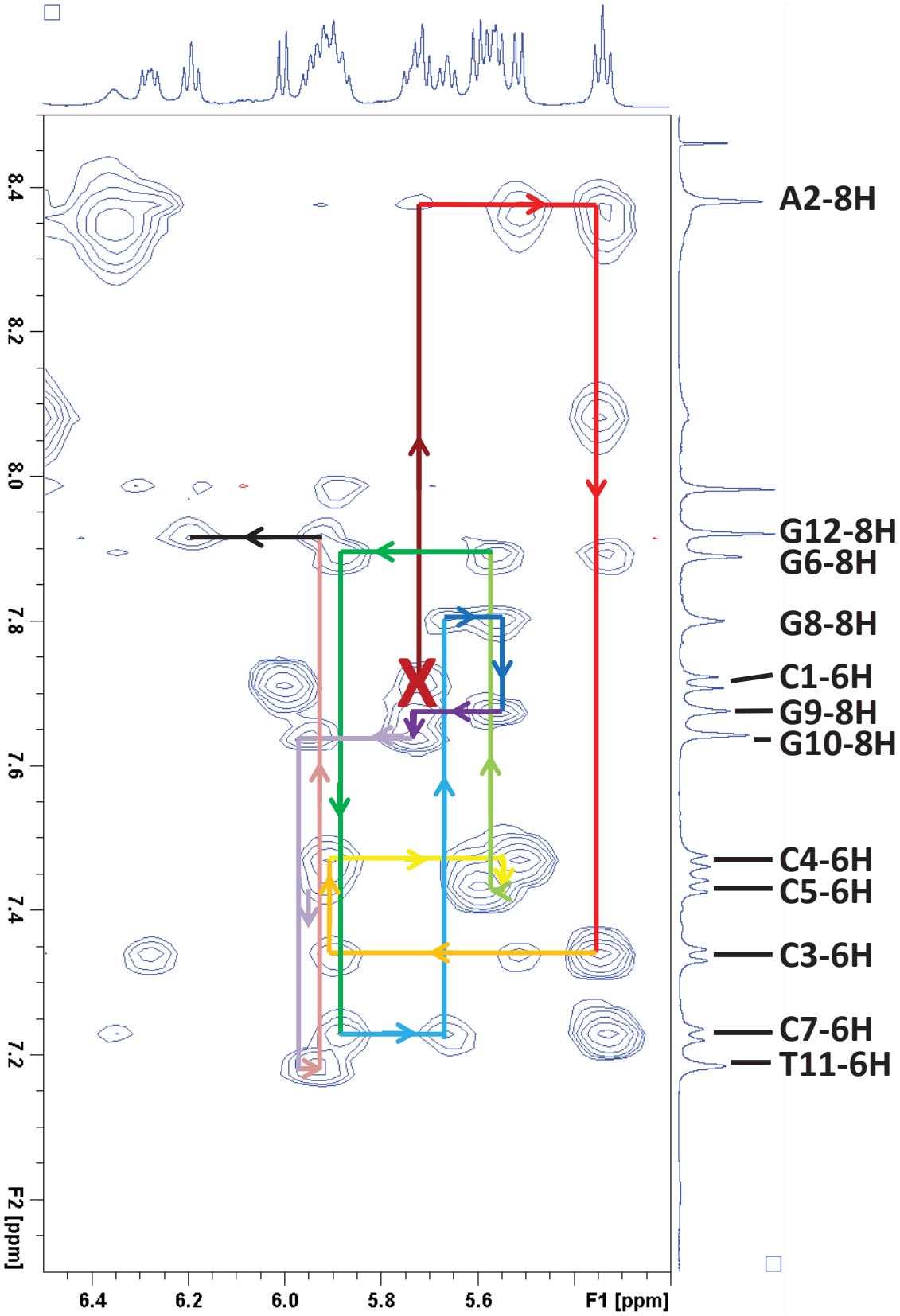
Sequence 122



Sequence 124



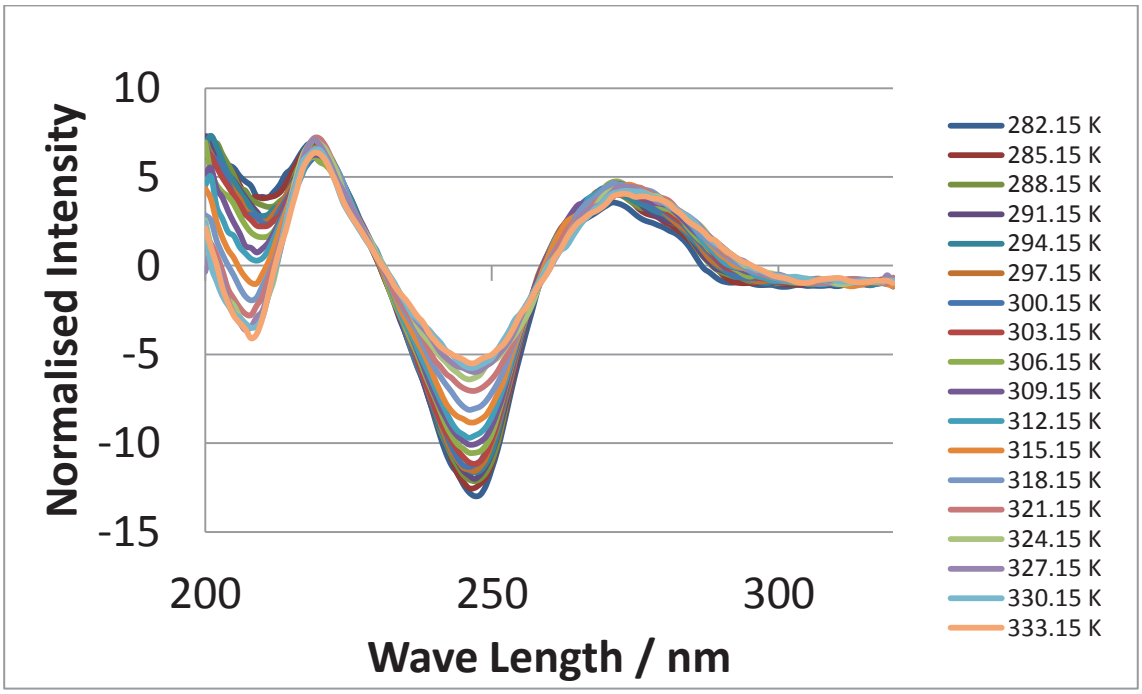
Sequence 125

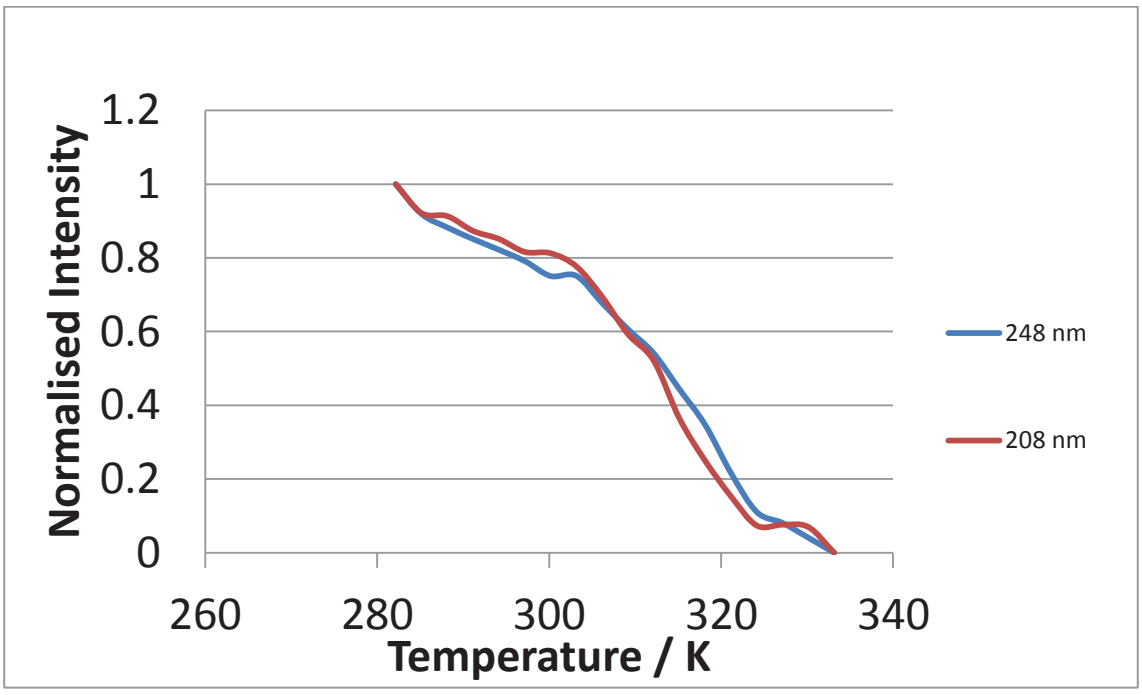


CD Spectroscopy Results

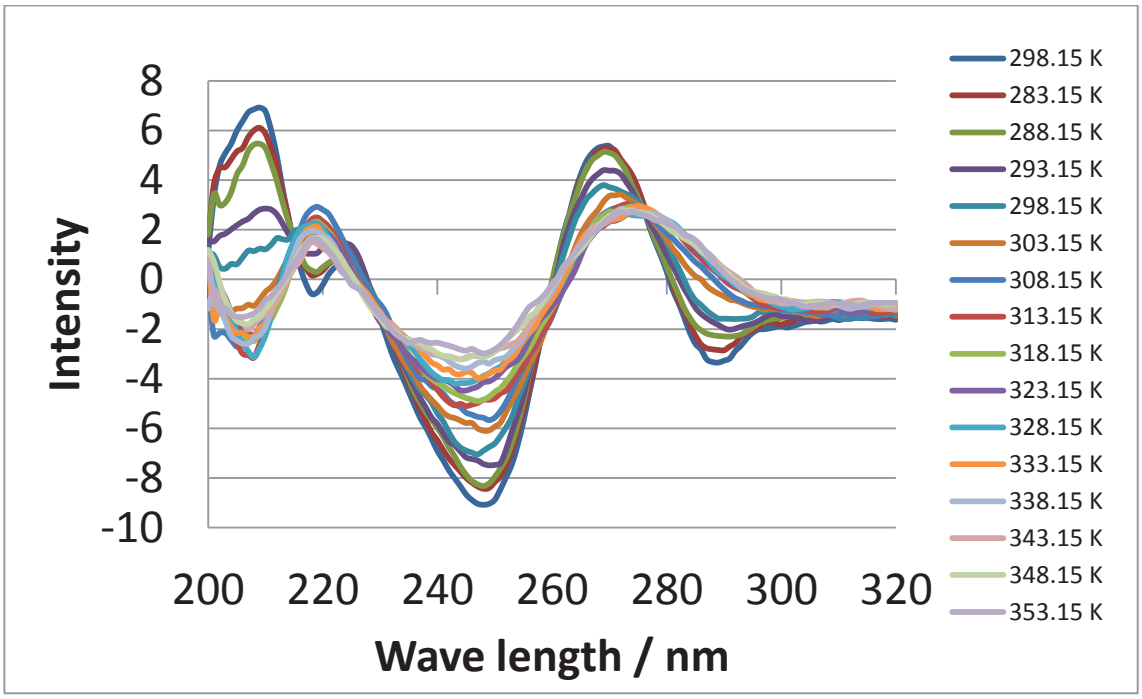
The CD spectroscopy melting spectra are given below along with the melting curves for each sequence.

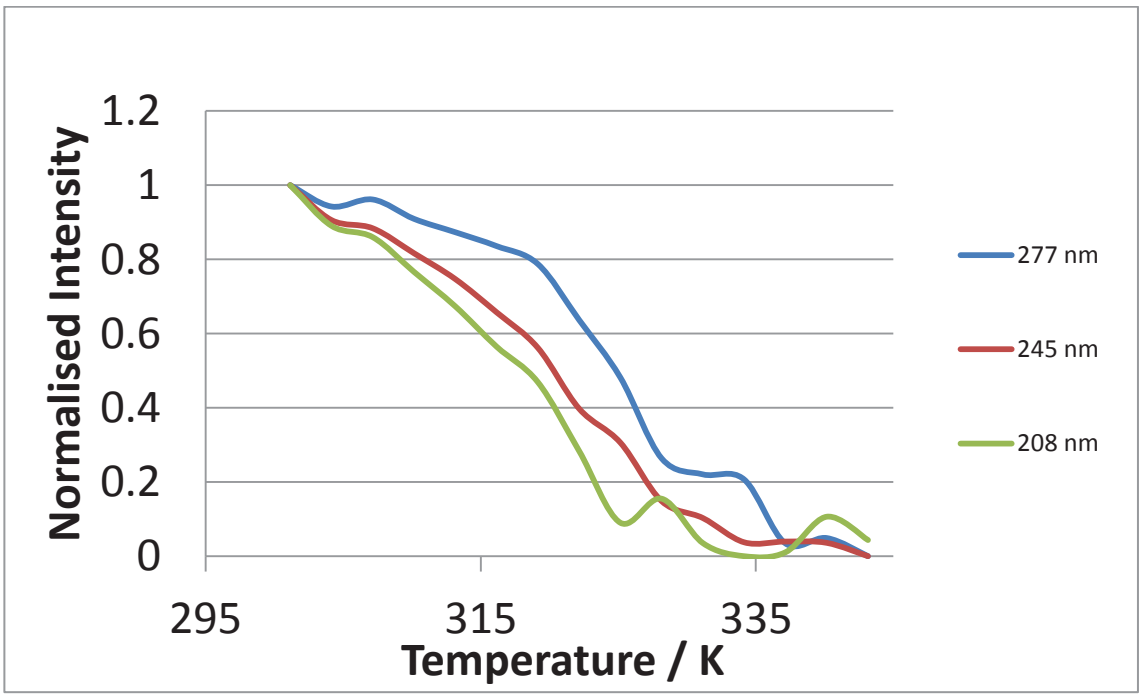
Sequence 121



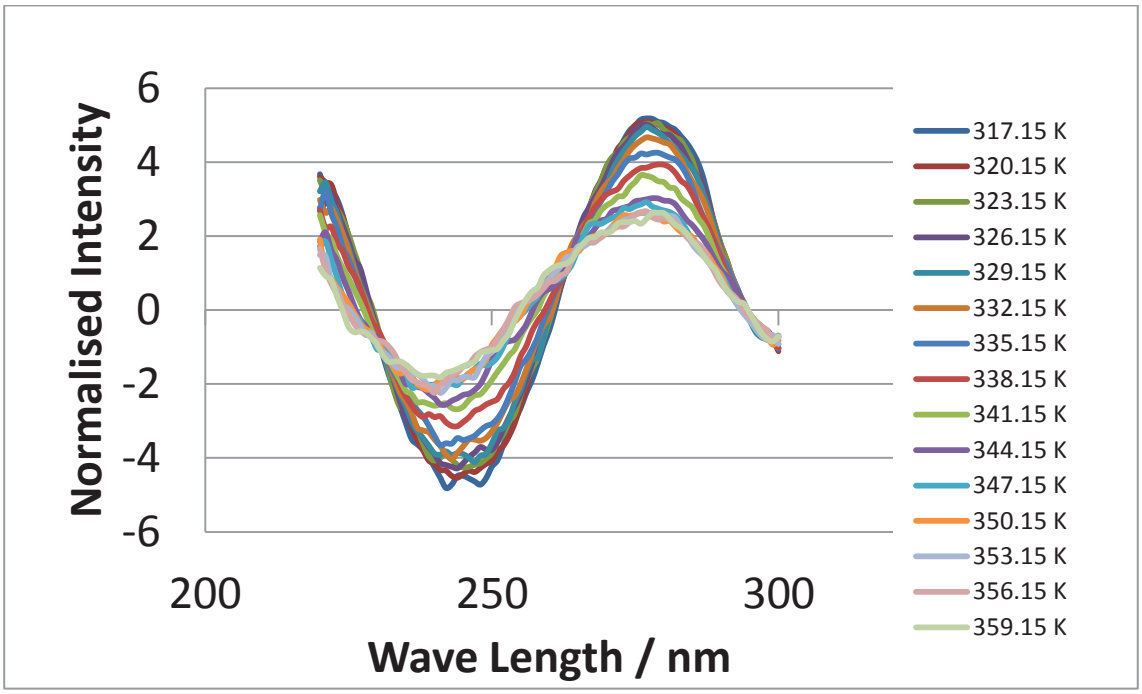


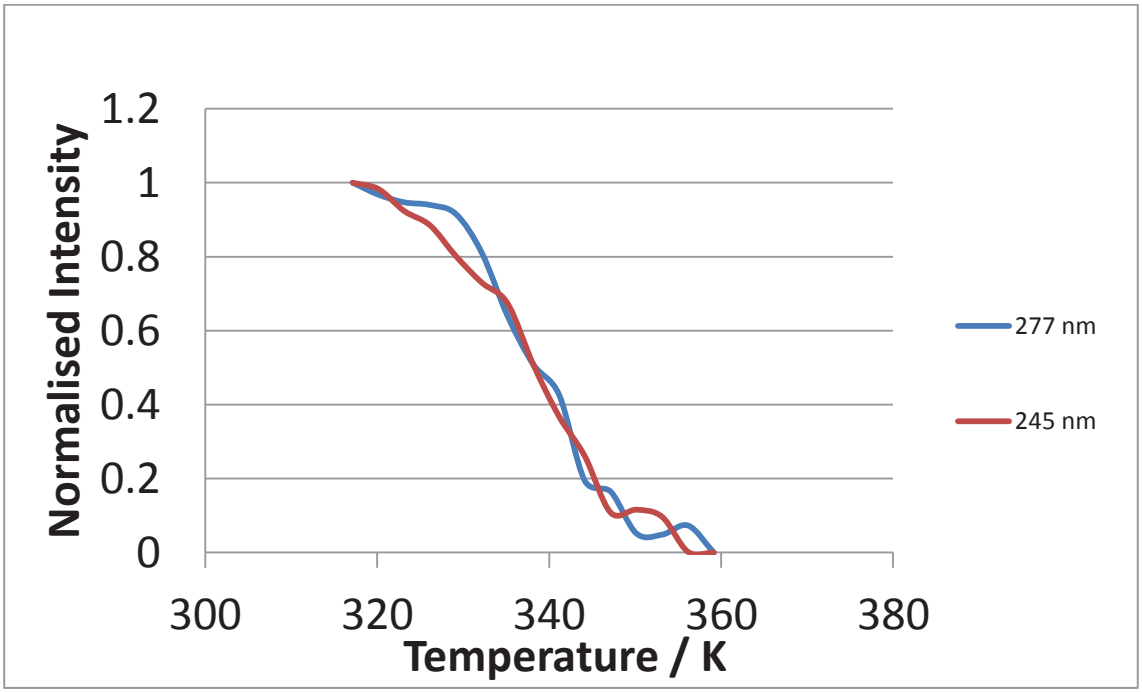
Sequence 122



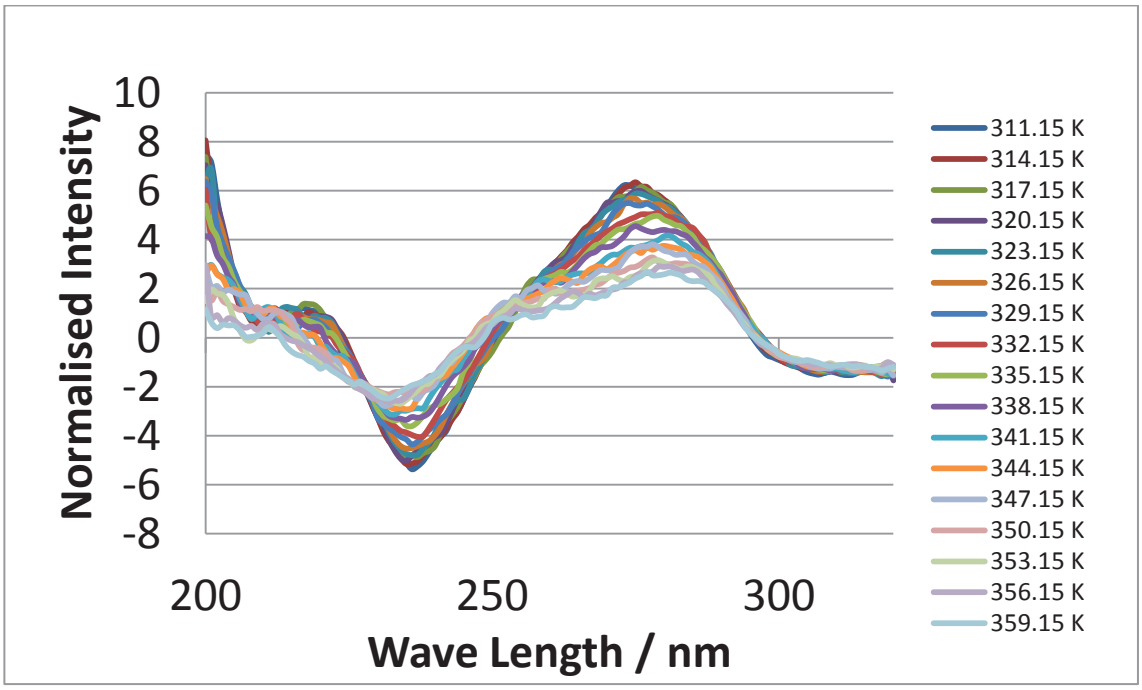


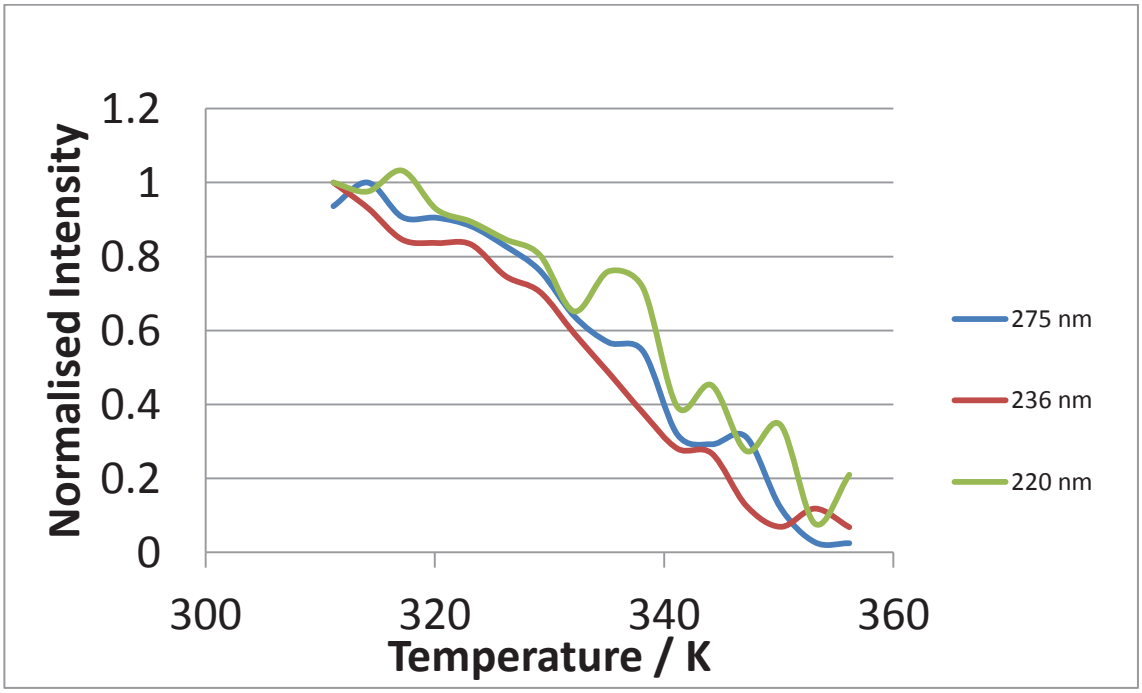
Sequence 123



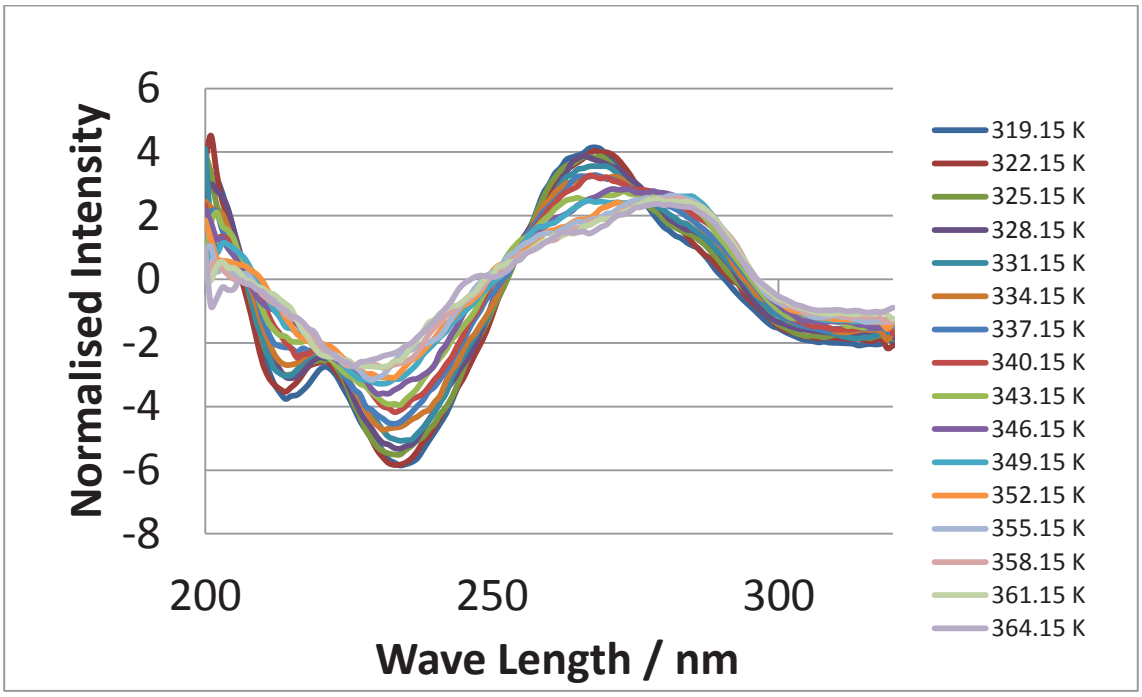


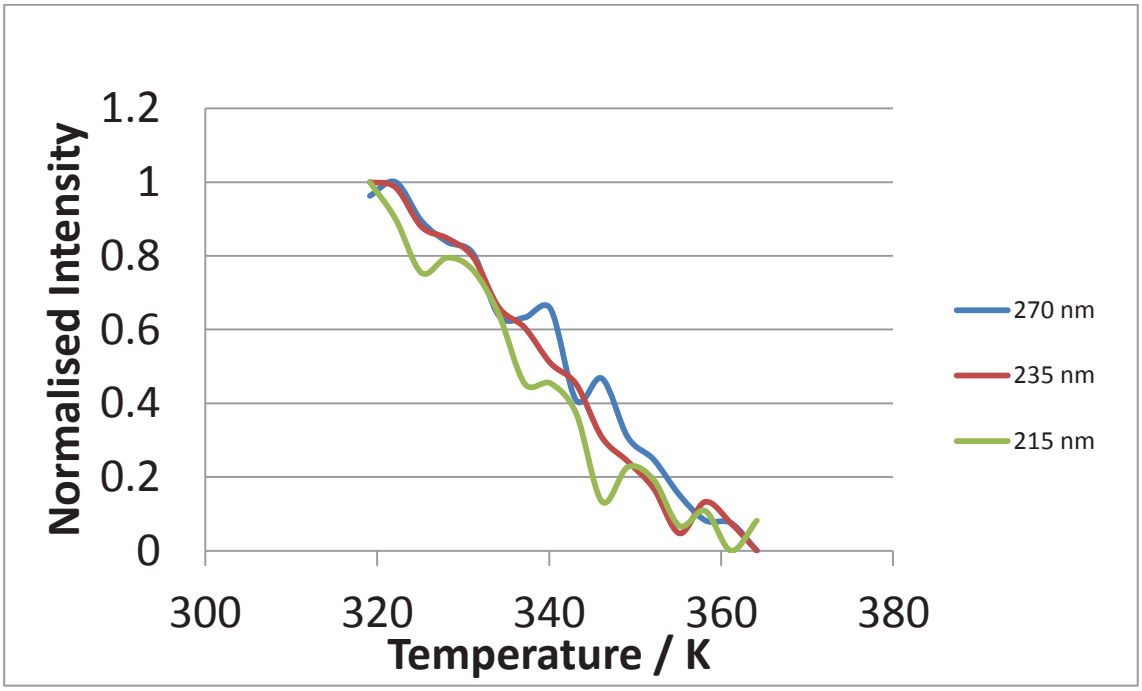
Sequence 124





Sequence 125





Melting Results

Melting points for base protons under high-pressure conditions for each dodecamer DNA sequence are given below. Melting points determined by fitting of an empirical function are indicated by a *, those which are determined visually are indicated by a ** and those which are rough estimates due to large gaps in the data are indicated by a ***. All other melting points have been determined by fitting with equation (23).

Sequence 121

Proton	0.1 MPa	50 MPa	100 MPa	150 MPa
C1-6H	311.837	313.402	315.621	316.1115
A2-8H*	311.28	311.382	312.008	312.1125
T3-6H*	313.971	312.695	315.146	314.938
T4-6H**	313	313.5	315	316
T5-6H*	311.617	314.262	315.146	312
A6-8H	314.746	315.873	315.325	317.463
T7-6H	314.86	314.725	314.602	314.6095
A8-8H	312.319	312.668	313.304	313.997
A9-8H*	315.252	316.216	316.375	317.8935
A10-8H	314.49	315.104	314.11	314.386
T11-6H***	314.476	313.056	315.778	316.5
G12-8H	312.885	314.049	314.386	314.379
Average	313.3944	313.911	314.7334	315.0325

Sequence 122

Proton	0.1 MPa	50 MPa	100 MPa	150 MPa	200 MPa
C1-6H	320.513	322.119	321.981	322.0317	321.883
A2-8H*	318.607	318.009	318.119	318.4433	319.477
T3-6H	322.111	321.465	321.621	321.348	320.81
T4-6H	321.186	321.93	321.956	322.2407	322.472
C5-6H	324.586	323.297	322.779	322.7937	323.278
T6-6H*	309.585	310.922	311.723	313.346	314.29
A7-8H*	311.84	311.831	311.778	312.2937	311.778
G8-8H	320.534	321.354	322.078	320.9177	319.599
A9-8H	320.642	321.515	321.781	322.7823	322.278
A10-8H***	324.985	325.307	326.001	327.2977	330.124

T11-6H* **	323.275	323.542	322.258	322.9713	324.718
G12-8H*	321.724	321.563	322.152	322.0623	324.29
Average	319.9657	320.2378	320.3523	320.7107	321.2498

Sequence 123

Proton	0.1 MPa	50 MPa	100 MPa	150 MPa	200 MPa
C1-6H	336.634	335.288	335.933	338.4153	340.616
A2-8H	337.034	335.437	336.195	338.4443	341.074
A3-8H**	339.331	338.983	339.812	341.5273	344.762
G4-8H	338.864	338.448	338.982	341.4983	344.266
T5-6H	338.333	337.781	338.07	339.6437	342.827
C6-6H* **	339.531	338.921	339.567	341.7383	343.517
G7-8H	338.938	337.281	337.731	340.193	341.895
A8-8H	338.486	337.366	337.628	340.0653	342.558
C9-6H* **	340.545	339.699	340.541	342.7737	343.705
T10-6H	338.96	337.947	338.421	340.741	343.897
T11-6H	337.528	336.622	337.413	339.1983	341.62
G12-8H	338.912	338.103	338.582	340.7783	343.253
Average	338.5913	337.6563	338.2396	340.4181	342.8325

Sequence 124

Proton	0.1 MPa	50 MPa	100 MPa	150 MPa
C1-6H	337.012	335.562	336.772	337.946
A2-8H* **	339.053	336.911	339.501	339.916
G3-8H	336.983	335.536	336.694	337.0685
G4-8H	339.788	338.047	341.08	341.1885
T5-6H	338.877	337.741	338.965	340.6225
C6-6H* **	341.52	339.031	341.781	341.273
G7-8H**	336	333	336.5	338
A8-8H	339.084	337.401	339.008	340.6895
C9-6H	338.439	338.121	339.702	340.2885
C10-6H* **	340.322	338.878	341.613	342.0805
T11-6H	337.309	335.034	337.839	338.443
G12-8H	338.466	336.273	338.511	338.9455
Average	337.9953	336.3017	338.3412	339.2436

Sequence 125

Proton	0.1 MPa	50 MPa	100 MPa	150 MPa	200 MPa
C1-6H	345.798	342.479	344.906	346.901	347.713
A2-8H* **	351.215	345.604	350.668	352.158	354.67

C3-6H	346.307	345.782	348.868	349.697	349.991
C4-6H	347.293	347.279	349.737	350.239	350.541
C5-6H* **	345.718	344.01	349.628	349.801	352.542
G6-8H	348.352	346.866	346.26	346.725	347.884
C7-6H	346.756	349.014	349.373	349.36	350.062
G8-8H	339.844	339.109	339.537	340.781	343.287
G9-8H	346.691	344.858	347.976	350.452	352.1
G10-8H	346.484	345.183	347.696	350.134	350.162
T11-6H	344.613	343.678	346.401	347.932	348.736
G12-8H	347.464	344.649	344.53	345.389	346.548
Average	345.9602	344.8897	346.5284	347.761	348.7024

Appendix Four – i-Motif Melting Data

The following table shows the results obtained for the melting of the i-motif, 5′-CCC TAA CCC TAA CCC TAA CCC-3′ at 0.1 MPa and various pH levels.

Pressure /MPa	Buffer	pH	NMR Melting Point /K
0.1	Phosphate	4.9	331.049
0.1	Phosphate	5.7	324.745
0.1	Phosphate	6.0	318.066
0.1	Phosphate	6.4	299.821
0.1	Citrate	4.1	326.187
0.1	Citrate	4.1	326.261
0.1	Citrate	4.6	332.046
0.1	Citrate	4.8	332.2
0.1	Citrate	5.0	329.111
0.1	Citrate	5.3	324
0.1	Citrate	5.5	316.902
0.1	Citrate	5.85	307.631
0.1	Citrate	6.2	301.038
0.1	Citrate	6.5	294

The following table shows the results obtained for the melting of the i-motif, 5′-CCC TAA CCC TAA CCC TAA CCC-3′ at 50 MPa, 100 MPa,150 MPa, and 200 MPa and various pH levels..

Pressure /MPa	Buffer	pH	Corrected pH	NMR Melting Point /K	Melting Point at 0.1 MPa at corrected pH /K
100	Phosphate	6.4	6.0	303.863	318
200	Phosphate	5.7	5	324.338	339
200	Phosphate	6.0	5.3	320.294	331
200	Phosphate	6.4	5.7	310.863	317.5
50	Citrate	5.85	5.75	309.831	310.8
100	Citrate	4.6	4.4	333.050	330
100	Citrate	5.5	5.3	318.39	322.4
100	Citrate	5.85	5.65	311.614	313.6
150	Citrate	5.85	5.55	312.674	315.6
200	Citrate	5.0	4.6	332.316	332
200	Citrate	5.5	5.1	321.023	327
200	Citrate	5.85	5.45	314.27	316.9
200	Citrate	6.2	5.8	306.095	310

Appendix Five –Final Concentrations of Amplified DNA

The following is a full list for all incubated DNA samples. Included are: the sample number, the sequencing reference number, the DNA type (SS, RFII or RFI), incubation temperature, incubation pressure, incubation time, concentration after amplification and 260 nm to 280 nm ratio after amplification. The concentration after amplification and 260 nm to 280 nm ratio were determined using a nanodrop spectrophotometer. Sample numbers 61 to 64 are reference samples which have remained frozen during sample gathering.

Sample number	Sequencing Reference Number	DNA type	Temperature /K	Pressure /MPa	Time /Hours	Concentration /ng μL^{-1}	260/280 Ratio
1	1	SS	283.2	0.1	99	13.1	2.1
2	2	SS	283.2	0.1	121	11.4	2
3	3	SS	283.2	0.1	171	15.4	1.93
10	4	SS	283.2	150	149	13.1	1.88
11	5	SS	283.2	150	173	15	2.11
12	6	SS	283.2	150	297	8	2.56
13	7	SS	283.2	150	221	17.3	1.89
14	8	SS	283.2	150	245	15	2.01
15	9	SS	283.2	150	269	12.7	1.95
16	10	SS	283.2	150	293	21.3	1.7
17	11	SS	283.2	150	317	13.6	1.94
18	49	SS	283.2	150	341	10.5	1.85
19	50	SS	313.2	0.1	91	22.6	1.83
20	51	SS	313.2	0.1	117	23.1	1.85
21	52	SS	313.2	0.1	140	21.5	1.94
22	53	SS	313.2	0.1	163	18.5	1.91
23	12	SS	313.2	50	92	60.5	1.85
24	13	SS	313.2	50	116	88.2	1.77
25	14	SS	313.2	50	140	54.5	1.9
26	15	SS	313.2	50	164	43.1	1.52
27	16	SS	313.2	100	92	10.3	2.12
28	17	SS	313.2	100	115	32.5	1.78
29	18	SS	313.2	100	139	14.2	1.77
30	54	SS	313.2	100	163	16.4	1.
31	19	SS	313.2	150	91	62.4	1.96
32	20	SS	313.2	150	155	76.7	196
33	21	SS	313.2	150	140	51.2	1.95
34	22	SS	333.2	150	163	64.8	2
35	23	SS	333.2	0.1	18	10.8	2.48
36	24	SS	333.2	0.1	20	15.4	2.13
37	25	SS	333.2	0.1	22	28.2	2.3

38	26	SS	333.2	0.1	24	57.8	1.95
39	27	SS	333.2	150	18	20.3	2.08
40	28	SS	333.2	150	20	59.2	2.04
42	29	SS	333.2	150	24	66.1	1.96
43	31	RFII	333.2	0.1	18	15.5	2.43
44	32	RFII	333.2	0.1	20	16.8	2.42
45	33	RFII	333.2	0.1	22	12.2	1.55
46	34	RFII	333.2	0.1	24	12.2	1.72
47	35	RFII	333.2	50	18	9.8	1.58
48	36	RFII	333.2	50	20	11.3	1.6
49	37	RFII	333.2	50	22	12.3	1.76
50	38	RFII	333.2	50	24	9.1	1.91
51	39	RFII	333.2	100	18	27.5	1.56
52	40	RFII	333.2	100	20	13.6	1.65
53	41	RFII	333.2	100	22	16.7	1.95
54	42	RFII	333.2	100	24	13.8	1.78
55	43	RFII	333.2	150	18	20.2	1.82
56	44	RFII	333.2	150	20	21.4	1.86
57	45	RFII	333.2	150	22	17.5	1.69
58	46	RFII	333.2	150	24	16.3	1.74
59	47	RFI	333.2	0.1	24	10.3	1.83
60	48	RFI	333.2	150	24	11.4	1.91
61	55	SS	263.2	0.1	REF	100	1.93
62	56	SS	263.2	0.1	REF	31.2	1.94
63	57	SS	263.2	0.1	REF	100	1.93
64	58	RFII	263.2	0.1	REF	100	1.86

Appendix Six – High-Pressure System Operational Manual

High-Pressure Operational Guide Lines

- **Pressurising Pump System**
- **High-Pressure NMR Cell**
- **High-Pressure Reaction Vessel**

WARNING!

This device has been designed for use at very high pressures and can be dangerous and potentially harmful to users and equipment. It is very important you read and understand these instructions before using this device. The maximum pressure the cell should be used at is 250 MPa. The maximum pressure of the pressure system and reaction vessel is 410 MPa. Use of the equipment above their rated pressure could result in failure.

Chapter 1 - Pressurising Pump System

1.1 Pump System Schematic

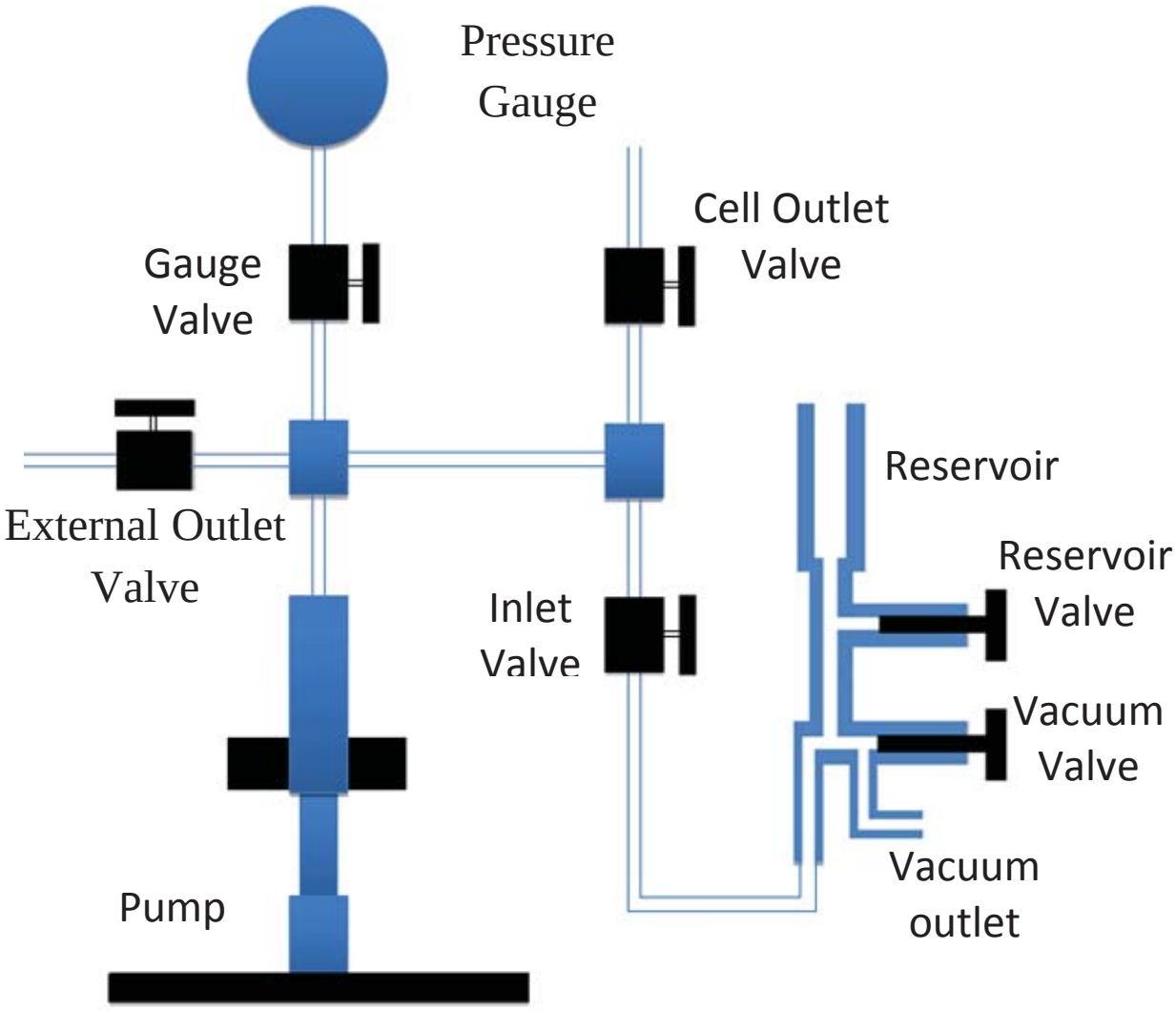


Figure 1.1: Pump system schematic

1.2 Safety Guidelines

- Ensure you are familiar with the Pump schematic before operating this system.
- Wear a face shield during pressurisation.
- Do not raise the pressure beyond the rating of equipment being used.

1.3 Priming the Pump System

For initial use the Pump system must be primed. If the pump system is already primed you may proceed to the next step. The pump system is primed with the use of a vacuum which draws in the hydraulic fluid; this is to be done prior to the cell assembly.

To prepare the system:

- Close off the gauge valve, the outlet valves and the reservoir valve.
- Open up the vacuum pump valve and the inlet valve.
- Wind the pump fully in.
- Attach the vacuum pump hose to the vacuum outlet.
- Fill the reservoir with the hydraulic fluid (It is important to shut off the gauge valve since the vacuum will damage the pressure gauge).
- Switch on the vacuum pump and run for 10 minutes.
- Close off the vacuum valve.
- Open up the reservoir valve slowly to allow the hydraulic fluid to be drawn into the pump system.

- Open up the gauge valve and detach the pump hose.
- Wind the pump all the way out to completely fill the pump chamber.
- Close off the inlet valve.

If necessary, more fluid can be drawn in by closing the outlet valve and opining the inlet valve and winding the pump back out. The reservoir valve should remain open.

Both outlet lines can be primed by opening the appropriate outlet valve and pumping fluid through the line until it flows out the end.

1.4 Hydraulic Fluid History

A detailed history of the hydraulic fluids that have been used by the system has to be kept. This is to ensure that future users have detailed information of what has been in the pump and to ensure that hydraulic fluids are periodically changed. When changing they hydraulic fluid please fill out the table in the back of this manual. Different fluids will have different acceptable lifetimes in the pump. Suggestions for some common fluids are in the technical details section. If your fluid is not on the list then it will be up to you to determine a change date.

1.5 Technical Details for Pump System

The following tables outline the pressure ratings and torque settings for the components of the pump system. More information can be found at www.highpressure.com.

1.5.1 Pressure Ratings

Part	Part number	Pressure rating (psi)
Pump	37-5.75-60	60 000
Gauge	6PG75	75 000
Valves	60-11HF4	60 000
Connections	60-23HF4	60 000
¼ inch high pressure tubing	60-HM4-2.75	60 000
⅛ inch high pressure tubing	60-9H2	60 000

1.5.2 Torque Settings for Glands

Part	Part number	Torque (ft lbs), (Nm)
¼ inch high pressure gland	60-2HM4	25 , 35
⅛ inch high pressure gland	60-2HM2	7.0 , 10
Adapter (for Gauge HF4 to XM6)	60-21HF4XM6	70 , 95
Adapter (for outlet HF2 to HM4)	60-21HF2HM4	25 , 35
⅛ inch high pressure gland to manifold valve	60-2HM2	7.0 , 10

1.5.3 Common Hydraulic Fluids and Change Times

Paraffin oil: 6 Months

Kerosene: 6 Months

Water: 3 Months

Chapter 2 - High-Pressure Cell

2.1 Safety Guidelines

DO:

- Exercise caution when pressurising the cell. The Perspex box should be used during preparation. This safety precaution is necessary to contain any fragments in the event of a tube fracture.
- Wear proper safety equipment such as a face shield when transferring the high-pressure NMR cell from the containment box to the NMR.
- Pre-test the cell at the target pressure for at least 15 minutes outside the NMR to assure integrity of the cell setup.
- Always align the centre of the cell tube serial number with the mark on the top of the cell base. Maintaining the same position of the tube relative to the manifold base assures consistent setup.
- Change the tube seal (TS01) after every use.
- Use care when inserting the cell into the magnet. Avoid striking the pressurised cell tube against objects.
- Leave the high pressure tubing tether attached to the cell between uses. Constant reseating of the fitting will reduce the seal integrity.

DO NOT:

- Pressurise the cell above 250 MPa (This is marked in red on the gauge).
- Insert a fully pressurised cell into the magnet. It is preferable, especially when working near the posted maximum, to insert the cell first then take it to pressure. It is recommended the cell pressure be below 100 MPa when inserting or removing the cell from the magnet.
- Over tighten the manifold base into the cell manifold. This can strip the threads or damage the seal surface and ruin the cell.
- Pressurise the cell while it is in the cell setup tool. The fit of the cell tube is very tight. If the setup is in some way improper, the tube may shift in the setup tool during pressurisation. A slight shift could fracture the tube.
- Use metal needles to fill the cell tube. The tube is made from a hard ceramic that can scrape metal off into the sample.

2.2 Safe Handling Apparatus

The high-pressure NMR cell is mounted on a specially designed safe handling apparatus. The apparatus is constructed from aluminium with stainless steel fittings. The following is a list of the key features of this design with references to their location on the schematic shown in Figure 2.1.

1. High-pressure cell. This is the position of the high-pressure cell at the base of the shaft.
2. Aluminium shaft. The high-pressure cell is attached to the shaft via a small PEAK plastic fitting. The length is such that, when lowered into the NMR spectrometer, the top of the shaft is sits on top of the shaft sleeve reducing

the weight of the shaft on the NMR probe. The high-pressure tether passes through the shaft and exits out the top.

3. Shaft sleeve. This sleeve allows the shaft to be positioned at the desired height and secured with the use of the locking nut located at the top of the sleeve. A fine adjustment collar is located on the outside of this sleeve so that its position within the alignment bracket can be changed.

4. Alignment bracket. This bracket has been developed so that the shaft can be easily and reproducibly positioned in an exact location over the top of the spin housing at the top of the NMR spectrometer. Three locking screws are positioned around this bracket so it can be secured to the spin housing in order to limit any movement of the high-pressure cell. When the cell is drawn up into the bracket prior to insertion into the NMR spectrometer, the cell is only visible from directly below the bracket. This shields the operator from any possible cell failure and makes it virtually impossible to strike the cell against objects.

5. Main arm. The main arm provides the mounting point for the high-pressure pump system. It has a full range of motion from the NMR spectrometer to flat up against the wall of the NMR lab.

6. Secondary arm. This extension to the main arm allows the alignment bracket to be manoeuvred around the nitrogen ports on top of the NMR spectrometer. It also contains two adjustment features which ensure that the shaft is positioned vertically.

7. High-pressure pump system mounting position.

8. Wall-mounting bracket

9. Locking arm. This arm can be attached to the main arm to lock it in position so the high-pressure pump can be operated without movement of the apparatus.

10. High-pressure tether. This tether transfers the pressure from the high-pressure pump system to the high-pressure cell. The tether has had a coil added to allow for the movement of the shaft up and down.

11.

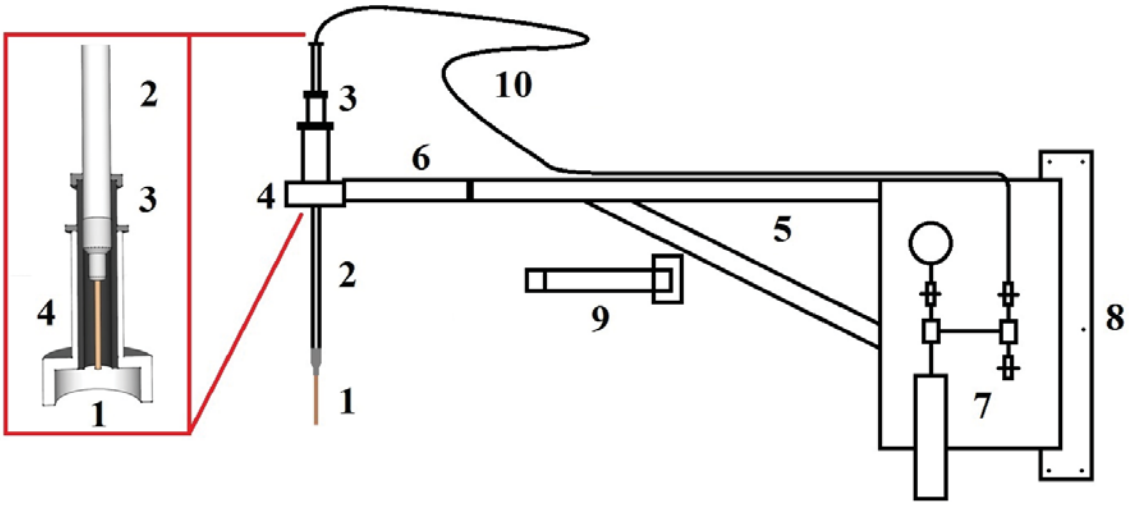


Figure 2.1: Safe-Handling Apparatus for the High-Pressure NMR Cell. This schematic outlines the important features of the safe-handling apparatus. Further information on these features is given in the text. (1) The high-pressure cell. (2) Aluminium shaft. (3) Shaft sleeve. (4) Alignment bracket. (5) Main arm. (6) Secondary arm. (7) High-pressure pump. (8) Wall-mounting bracket. (9) Locking arm. (10) High-pressure tether.

2.3 About the High-Pressure Cell

The pump system and high-pressure tether must be primed before the cell is assembled (see “Priming the Pump System”). If the pump system has not been used in a long time the hydraulic fluid may need to be changed (see “Hydraulic Fluid History”).

The cell was designed so that the cell tube can be easily removed and cleaned if necessary. As such the seal (TS01) is single use. An ancillary component called the tube seat (TCSN-M) serves as a cushion between the ceramic tube and metal surface of the manifold base. Figure 2.1 shows the relative position of the tube seal (TS01) and tube seat (TCSN-M).

The tube seat is a permanent component and should not be removed from its position in the bottom of the manifold base. If it dislodges it should be reinserted according to the diagram with the internal bevelled edge of the tube seat facing the top of the tube. The tube can then be used to reposition the tube seat.

The primary seal is provided by the part TS01. The seal is single use only. For setup it should be placed on the end of the manifold valve section piece. A small amount of hydraulic fluid should be pumped through the end of the valve section to minimise any air gap.

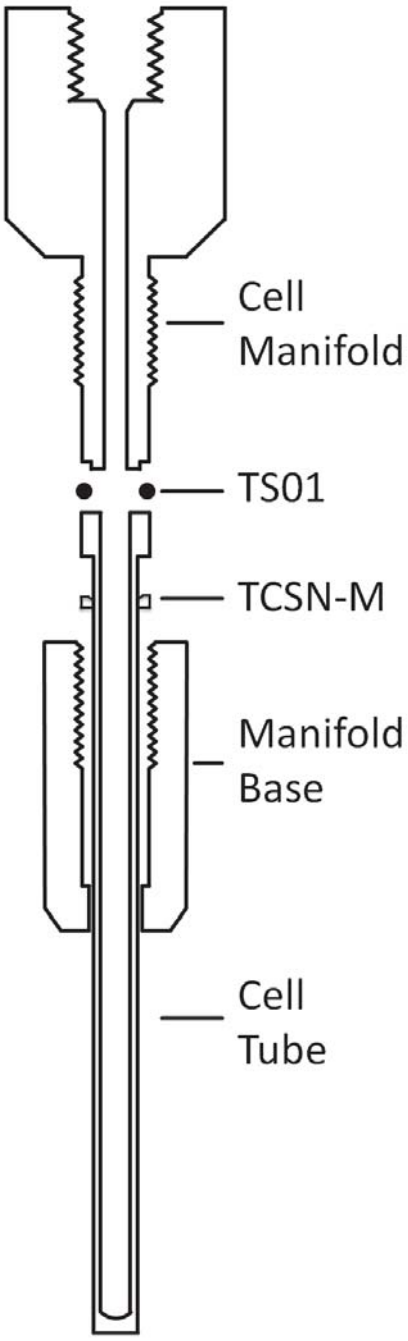


Figure 2.2: High pressure cell schematic

For proper positioning of the tube the Cell Setup Tool should be used. This tool helps keep the tube axially aligned with the NMR cell. Improper positioning of the tube can prevent the cell from inserting into the NMR.

2.4 Preparation of the High-Pressure Cell

Due to the high pressures involved with this equipment it is important that you take care while moving and applying pressure to the cell.

If operating the cell at raised temperatures, reduce the cell pressure by 5 MPa before starting to heat in the cell. Leave the sample for 15 minutes at the target temperature to allow it to equilibrate and then raise the pressure back to the operating pressure. If operating near the maximum, continue to monitor the pressure reading during the experiment.

The following procedure outlines the cell filling and assembly procedures.

2.4.1 Preparation of Pump System

- If there is no hydraulic fluid in the pump system it will need to be primed. See “Priming the Pump System”.
- Ensure that the inlet and external outlet valves are closed and that the gauge and the cell outlet valves are open.
- Place a fresh TCSN01 seal on the end of the cell manifold and pump a small amount of hydraulic fluid through the end of the cell manifold. Close the cell outlet valve.

2.4.2 Insertion Sample in the Cell

- Fill the tube to the top using a **glass pipette**. Ensure there are air no bubbles.
- Carefully withdraw the top 3 cm of the sample from the tube. A marked pipette is recommended.
- Slowly pipette in the hydraulic fluid in from the top of the tube. When done carefully, the fluid will slide down the inside edge of the tube and fill up without air bubbles. Fill tube all the way to the top.

2.4.3 Cell Assembly

- Thread the cell tube in to the manifold base. Ensure that the centre of the serial number on the tube lines up with the indentation on the top of the base.
- Carefully screw the manifold base on to the cell manifold until hand tight.
- Lower the cell into the cell alignment tool.
- Rotate the cell alignment tool several times.
- Using the torque wrench and spanner provided, tighten the manifold base a small amount.
- Repeat the above two steps until the desired talk is achieved.
- Remove the cell from the alignment tool and wipe away any excess oil.
- The cell outlet valve can now be opened.

2.4.4 Torque Settings for Cell Assembly

Typical: 130 inch lbs. Max: 160 inch lbs.

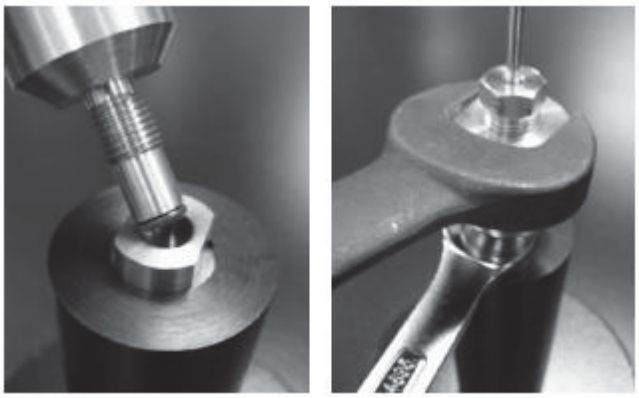


Figure 2.3 Cell assembly in the cell setup tool

2.5 Operation Safe Handling Apparatus and Pump System

The following procedure outlines the steps needed to test the cell, insert the cell into the NMR and to remove it from the NMR.

2.5.1 Pressurising the Cell

- Carefully position the cell in the clear Perspex box up on the raised platform.
- Ensure the locking arm is in place.
- Increase the pressure to the desired level and leave for 15 minutes. This will assure the integrity of the cell setup.

- Reduce the pressure to less than 100 MPa.

2.5.2 Cell Transfer to the NMR

- Raise the shaft up until the marked line is just above the top of the fine adjustment sleeve and tighten the shaft lock.
- Check to see that the cell tube is safely within the confines of the alignment bracket as shown (Figure 2.6). If the cell is too low it can be raised further by the fine adjustment collar. **Failure to do this can lead to the cell being damaged**.
- Remove the locking arm.
- Position the alignment bracket over top of the NMR.
- Slowly and carefully lower the shaft into the NMR. Do not force the shaft down.
- Evenly tighten screws on the alignment bracket.
- The cell can now be brought back up to the operating pressure.

2.5.3 Cell Transfer from the NMR

- Reduce the pressure in the cell to less than 100 MPa.
- Raise the shaft up until the marked line is just above the top of the fine adjustment sleeve and tighten the shaft lock.
- Check to see that the cell tube is safely within the confines of the alignment bracket and rise if necessary.

- Loosen the screws on the alignment bracket.
- Return the safe handling apparatus to its original position and secure with the locking arm.
- Carefully lower the shaft into the clear Perspex box.
- Reduce the pressure in the cell to zero.
- Open the inlet valve to release any excess pressure in the system.
- Close the cell outlet valve.

The cell can now be disassembled.

2.6 Cleaning the Cell

For extensive cleaning the NMR tube can be exposed to detergents suitable for quartz cuvettes. If nitric acid is to be used it is recommended that only dilute solutions be used and for short exposure times. The NMR tube can be autoclaved and subjected to temperatures up to 200 ºC.

Chapter 3 - High-Pressure Reaction Vessel

3.1 Safety Guidelines

- Never open the isolation valve when there is a pressure differential across the valve.
- Wear a face shield when disconnecting the high pressure connection.
- Avoid unnecessary bending of the high-pressure connection.
- Make sure that the high-pressure connection is secured properly.
- Make sure to minimise trapped air.
- Always secure the reaction vessel to the base plate.
- The reaction vessel is ferromagnetic and should never be taken within the marked areas within the NMR Lab.
- Avoid striking the high-pressure line against any objects.
- Use caution when moving the reaction vessel due to its weight.

3.2 About the High Pressure Reaction Vessel

The Series “R” O-ring Closure Reactor is easily assembled and disassembled with minimal torque required for complete engagement. The reaction vessel is made from Type 1150 stainless steel. Sealing is accomplished by a highly reliable combination of O-ring and separate metal back-up ring. The wedge shaped back-up ring is designed to expand and contract as pressure increases or decreases. Consequently, the O-ring is continuously confined with no clearance for extrusion.

Minimal initial torque is required to effect a positive seal. The operating temperature on this vessel must be restricted to 120°C maximum, due to the BUNA-N (nitrile) O-ring.

The reaction vessel is housed in a custom frame with a built in isolation valve so it can be easily disconnected and transferred to a water bath for incubation while still under pressure.

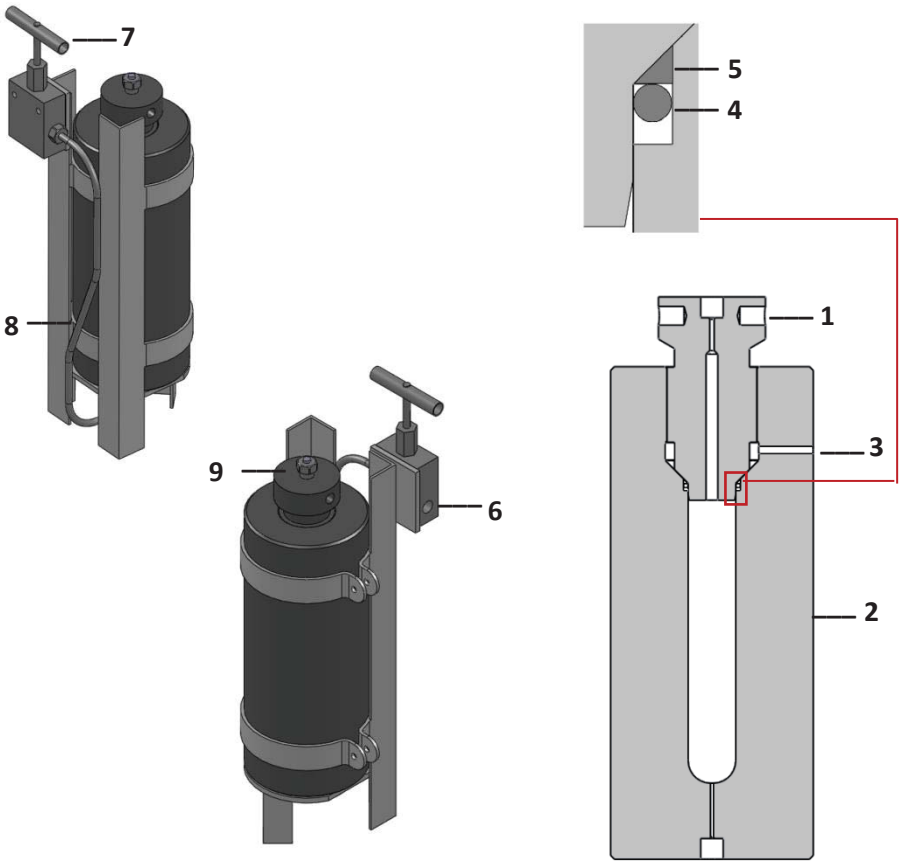


Figure 3.1: Reaction Vessel. (1), Cover. (2), Body. (3), Vent hole. (4), "O" ring. (5), Back up ring. (6), High-pressure tether connection point. (7), Isolation valve. (8), Pressure Line. (9), Plug.

3.3 Preparation and Loading of the Reaction Vessel

The following instructions outline steps to prep and load the reaction vessel. Preparation of the reaction vessel will only have to be performed when there is air in the high pressure line. To avoid this, always keep the isolation valve closed when not connected to the pump system.

3.3.1 Preparation

- Attach the reaction vessel to the pump system (see 3.4.1: Connecting to pressure system).
- Open the isolation valve.
- Pump through several millilitres of hydraulic fluid to ensure there is no air trapped in the line.
- Close the isolation valve.
- Disconnect the reaction vessel from the pump system.

3.3.2 Loading the Reaction Vessel

- Place capillary holder in the reaction vessel.
- Fill the reaction vessel with hydraulic fluid up the level of the “O” ring.
- Screw on the cap and tighten with the bar.

3.4 Operation of the Reaction Vessel

3.4.1 Connecting to Pressure System

- Close the inlet and the cell outlet valves and open the external outlet and gauge valves
- Pump a small amount of hydraulic fluid out the end of the high-pressure connection.
- Secure the reaction vessel to the base plate and attach the high-pressure connection to the isolation valve.
- Use the torque wrench to tighten the connection to 7 Nm
- Open the isolation valve

3.4.2 Pressurising and Disconnecting Under Pressure

- Insure that the isolation, external outlet and gauge valves are open and that the inlet and cell outlet valves are closed.
- The system can now be pressurised (make note of the pressure within the reaction vessel).
- Close the isolation valve.
- Reduce the pressure in the system to zero.
- Open the inlet valve to release any excess pressure.
- Carefully disconnect the high pressure connection from the isolation valve.

3.4.3 Connecting Under Pressure and Depressurisation

- Connect the reaction vessel to the pressure system keeping the isolation valve **closed.**
- Close the inlet and the cell outlet valves and open the external outlet and gauge valves.
- Raise the pressure of the pump system to equal the pressure within the reaction vessel.
- The isolation valve can now be opened.
- Reduce the pressure to zero.
- Open the inlet valve to release any excess pressure.
- Carefully disconnect the high pressure connection from the isolation valve.
- Close the isolation valve
- The reaction vessel can now be opened

Chapter 4 - Further Information

If you are unsure about anything not outlined in this manual consult with the manufacturer.

High-pressure NMR cell.

www.daedalusinnovations.com

High-pressure pump and fittings and reaction vessel.

www.highpressure.com

4.1 Hydraulic fluid History

Fluid Used	Date	Date to change	Comments

Appendix Seven – Digital Appendices

- Cytosine/Cytidine Hydrolysis Results
- DNA Dodecamer Melting Results
    - CD data
    - Fitting data
    - Melting data
- DNA Hexamer Melting Results
    - Signal tracking
    - CD data
    - Hexamer melting results
- High Pressure Pump System
    - 3D model
    - High-pressure equipment information
    - Manuals
- i-Motif Melting Results
    - CD results
    - I-Motif NMR melting
- NMR Data
    - Cytosine hydrolysis
    - Dodecamer melting
    - Hexamer melting
    - I-Motif melting