An NMR Study of a Natural and a 3′-*S*-phosphorothiolate Modified DNA Triplex

Kathryn Evans

Submitted in accordance with the requirements for the degree of
Doctor of Philosophy

The University of Leeds
School of Chemistry

September 2012

The candidate confirms that the work submitted is her own and that appropriate credit

has been given where reference has been made to the work of others.

Acknowledgements

First and foremost I would like to thank my supervisor, Dr. Julie Fisher for all of her help, support and guidance throughout the duration of my PhD. I could not have wished for a better supervisor and have enjoyed every moment that I have worked for her.

I would also like to thank Prof. Rick Cosstick, Dr. Inder Bhamra and Alannah Clelland from the University of Liverpool for synthesising and providing the $3'$-S-phosphorothiolate modified oligonucleotide which was necessary for the completion of this project. I would also like to acknowledge Dr. Richard Wheelhouse from the University of Bradford for his help and assistance with the UV thermal melting studies. My thanks also go to Simon Barrett, Arnout Kalverda and Gary Thompson for providing both NMR time and knowledge and to the technical staff in the School of Chemistry for their support.

On a more personal note I would like to thank the Fisher Group, both past and present for their advice, support and company throughout my time at the University of Leeds. Finally I must thank my friends and family, and in particular my parents, who have provided constant love and support throughout this often difficult time.

Abstract

Triple helical DNA is formed when purine or pyrimidine bases of a triplex forming oligonucleotide (TFO) occupy the major groove of a homopurine-homopyrimidine double helix, and interact with it via Hoogsteen hydrogen bonding. There is significant interest in triplex structures due to their prevalence in biology as they are the target complex for antigene therapeutic approaches.

Triplex structures which comprise a DNA duplex and RNA TFO have previously been found to be more stable than the equivalent all DNA systems. There are synthesis and stability issues surrounding the use of RNA-based TFOs, hence there is interest in developing chemically-modified TFOs which show at least the same duplex-binding affinity. In this project, a 3′-*S*-phosphorothiolate linkage is the modification of interest, which has previously been shown to alter the conformation of deoxyribonucleic acid systems to that of ribonucleic acids.

A 12 base pair homopurine-homopyrimidine hairpin duplex that is a target for a TFO was characterised by NMR spectroscopy and high resolution structures were generated. The hairpin was shown to adopt a B-type helix with predominantly south sugar puckers.

Similar analysis was performed on a native triplex comprising the hairpin and a pyrimidine TFO. Binding of the third strand was found to cause little structural changes to the hairpin. Two non-adjacent 3′-*S*-phosphorothiolate modifications were then incorporated into the TFO and the effects determined by NMR analysis. The result was very localised conformational changes in the deoxyribose sugars attached to the modifications, as well as those 3′ to it, from a DNA-like south to an RNA-like north pucker.

Finally, UV thermal melting analysis of the target hairpin and triplexes was performed. The stability of the hairpin was found to be pH independent, whereas the triplex structures were only stable at acidic pH. The 3′-*S*-phosphorothiolate modification was found to stabilise the triplex by up to around 7 °C and to increase the pH range for triplex formation. It is thought that the increased stability is due to favourable base stacking interactions as a result of the modification.

Table of Contents

List of Figures

List of Tables

Abbreviations

δ	chemical shift
Δδ	chemical shift difference
τ_m	mixing time
°	degree
1D	one dimensional
2D	two dimensional
3D	three dimensional
3′-S	3′-S-phosphorothiolate
A	adenosine
Å	angstrom
C	cytidine
CNS_solve	Crystallography and NMR Systems solve
CORMA	complete relaxation matrix analysis
CPG	controlled pore glass
CPT	camptothecin
DG	distance geometry
DMT	dimethoxytrityl
DNA	deoxyribonucleic acid
DQF-COSY	double quantum filtered correlation spectroscopy
DSC	differential scanning calorimetry
dU	deoxyuridine
G	guanosine
HIV	human immunodeficiency virus
HPLC	high performance liquid chromatography

IRMA	iterative relaxation matrix approach
ISPA	isolated-spin-pair approximation
LNA	locked nucleic acid
MARDIGRAS	Matrix Analysis of Relaxation for Discerning the Geometry of an Aqueous Structure
MD	molecular dynamics
mRNA	messenger ribonucleic acid
nm	nanometre
NMR	nuclear magnetic resonance
nOe	nuclear Overhauser effect
NOESY	nuclear Overhauser effect spectroscopy
PCYT	protonated cytidine residue
PDB	protein data bank
PNA	peptide nucleic acid
ppm	parts per million
rMD	restrained molecular dynamics
RMSD	root mean square deviation
RNA	ribonucleic acid
STHY	3′-S-phosphorothiolate modified thymidine residue
T	thymidine
T_m	melting temperature
TFO	triplex forming oligonucleotide
TOCSY	total correlation spectroscopy
Topo I	Topoisomerase I
U	uracil
UV	ultraviolet
W-C	Watson-Crick

Chapter 1

Introduction

Chapter 1.   Introduction

1.1 Nucleic acids

Nucleic acids have many functions in living organisms. They are responsible for storage and interpretation of genetic information, catalyse many chemical reactions and are the active elements of many viruses. Nucleic acids are considered to be among the most important molecules in biology, and the study of their structure and function is fundamental to an understanding of life[1]. The first insight to the function of DNA as the molecule of genetic material was reported in 1944 by Avery *et al.*[2], who discovered that DNA from one bacterial species was able to inherently alter the characteristics of another species and that genetic information could be passed down through generations.

In 1953, Watson and Crick proposed that DNA forms a double helical structure consisting of two polynucleotide strands connected by hydrogen bonding between complementary bases[3]. Franklin and Gosling not only confirmed this structure using X-ray crystallographic methods, but also discovered that DNA can adopt different types of helical structures[4]. These discoveries were fundamental to the understanding of nucleic acid structures that we have today.

1.2 Structure of nucleic acids

Nucleotides are the 'building blocks' of both ribonucleic acid (RNA) and deoxyribonucleic acid (DNA). Nucleotides are constructed from three components; a heterocyclic purine or pyrimidine base, a pentose sugar and a phosphodiester group.

In DNA, the sugar is deoxyribose whereas in RNA the sugar is ribose. The purine

bases are adenine (A) and guanine (G), and the pyrimidine bases are cytosine (C),

thymine (T) in DNA and uracil (U) in RNA[1] (Figure 1.1).

deoxyribose sugar   **ribose sugar**

heterocyclic bases
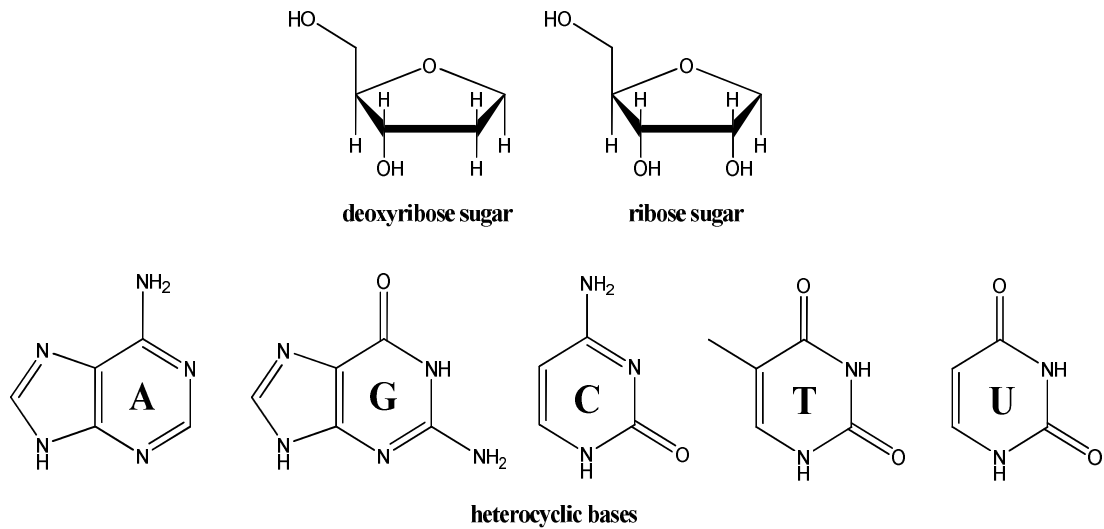
Figure 1.1. Sugars and bases which make up DNA and RNA nucleotides.

The pentose sugars and phosphodiester groups in a nucleic acid are the major

structural component. Alternating sugars and phosphates make up the backbone of

the polymer, with the phosphate linking the 3′-hydroxyl of one sugar to the 5′-

hydroxyl of another. Nucleic acids therefore have direction, with the two ends of the

molecule termed the 3′ and 5′ termini as indicated in Figure 1.2.

Chain direction

5′ to 3′

5′ end

n-1

n

n+1

3′ end
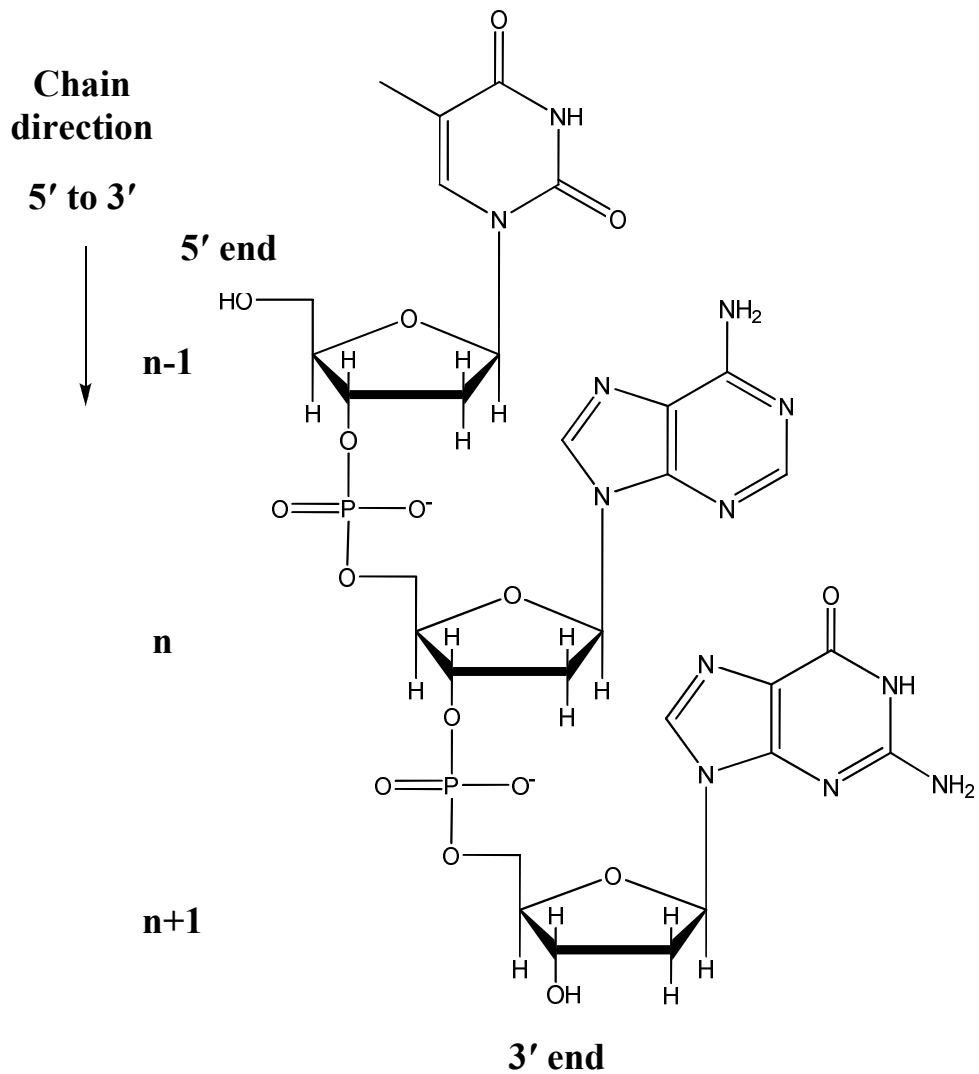
Figure 1.2. Molecular structure of DNA with the chain direction highlighted.

The internal core of the double helix consists of the heterocyclic bases, which participate in hydrogen bonding with a complementary base on an adjacent strand. This is termed Watson-Crick base pairing and is specific such that adenine pairs with thymine (or uracil in RNA molecules) and cytosine pairs with guanine. C-G base pairs are held together through three hydrogen bonds whereas A-T (or A-U) base pairs contain only two, as outlined in Figure 1.3.
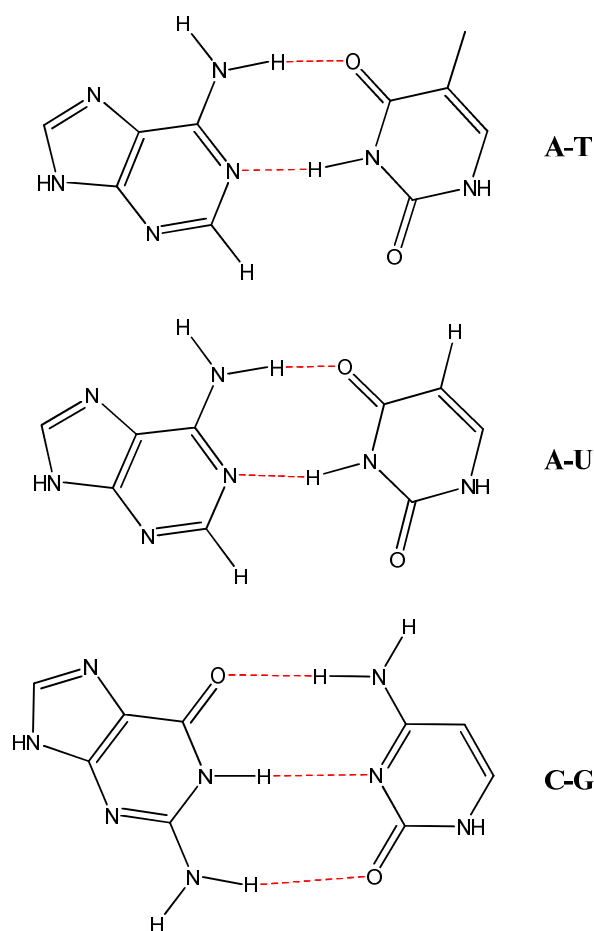
Figure 1.3. Watson-crick base pairing between complementary bases. Hydrogen bonds are highlighted in red.

These hydrogen bonding interactions between bases contribute to the formation of a stable duplex. Other duplex stabilising factors include interactions between the solvent and sugar-phosphate backbone, and vertical base stacking between adjacent base pairs[5]. There are several forces which are involved in base stacking; namely dipoles, π-π interactions, dipole-induced dipole moments and London dispersion forces (induced-dipole induced-dipole interactions)[1]. The nucleobases possess a permanent dipole which can induce a dipole in the π-electron system. In addition, fluctuations within the electronic charge distribution of the system can give rise to an instantaneous dipole, which then induces a temporary dipole on a neighbouring atom

or molecule resulting in an attractive force between the two[1]. Hydrophobic interactions that are associated with base pairing and stacking also have a role in stabilising the structure, as formation of hydrogen bonds between bases and vertical stacking of bases eliminates water molecules from the duplex, resulting in a favourable entropy gain for the system[6].

The deoxyribose or ribose sugars in nucleic acids are not planar - which would result in unfavourable interactions between substituents on the ring, but instead are puckered with one or more ring atoms being out-of-plane[7]. Sugars can adopt one of two conformations, and DNA sugars have been found to preferentially adopt the C2′-*endo* or south pucker, whereas RNA sugars tend to reside in the C3′-*endo* or north conformation[1, 8]. Figure 1.4 illustrates the equilibrium between the two possible conformational extremes.

North
C3'-*endo*
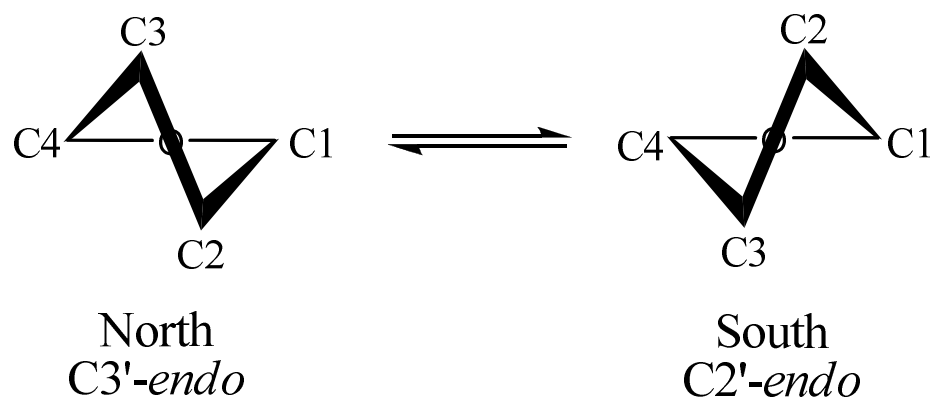
South
C2'-*endo*

Figure 1.4. North and south sugar puckers that are typically adopted by RNA and DNA sugars, respectively.

The position of equilibrium depends on two stereoelectronic effects. The anomeric effect occurs due to favourable overlap between the lone pair on the ring oxygen and

the σ* molecular orbital of the C1′-N bond. This effect is maximised in the north

pucker as the C1′-N bond is best aligned for anomeric overlap (Figure 1.5).

Figure 1.5. Orbital positions in the south and north sugar puckers, showing anomeric overlap potential in North pucker.

The *gauche* effect, which describes the inclination of two electronegative substituents

on adjacent carbons to adopt a gauche rather than a trans orientation (*i.e.* to have a

torsion angle of less than ± 90 °), is the second principle which affects the preferred

conformation of the sugar ring. When there is no 2′ OH substituent, as is the case for

DNA sugars, the *gauche* effect overrides the anomeric effect so that the 3′ oxygen

can adopt a gauche position relative to the 4′ C-O bond as shown in Figure 1.6 (a).

The additional 2′ OH substituent in RNA sugars causes a switch to the north pucker

governed by the anomeric effect and the *gauche* effect between the 2′ OH and the 4′

C-O bond, which is highlighted in Figure 1.6 (b).
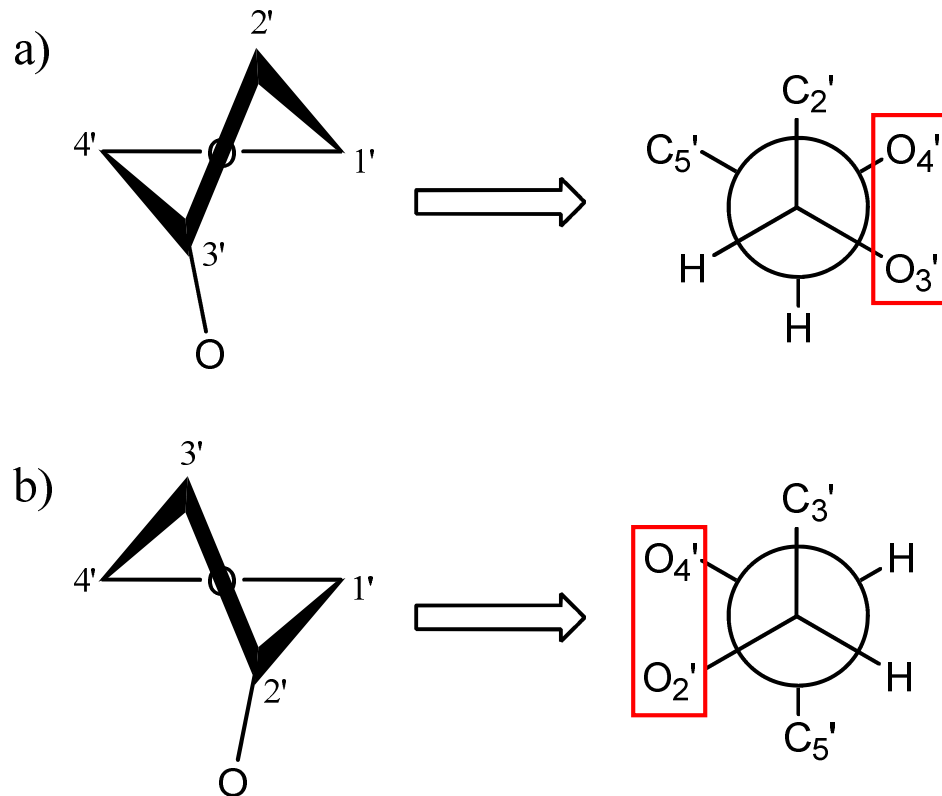
Figure 1.6. Gauche positioning of electronegative substituents in a) deoxyribose sugars and b) ribose sugars.

Nucleic acids can form three main types of double helix, termed A-, B- and Z-type, with the major difference between these being caused by the conformational preference of the sugar ring[9]. An A-type helix arises when the nucleotides adopt essentially north sugar puckers and so is typical of RNA. These are a shorter, more compact helix than B-type structures, which are characteristic of DNA and contain mainly south sugar puckers (Figure 1.5). Z-type helices result from alternating north and south puckers, and are generally unfavourable except under conditions such as an alternating purine-pyrimidine sequence or high salt content[8]. A- and B-type structures are right handed, whereas Z-DNA presents a left handed double helix.
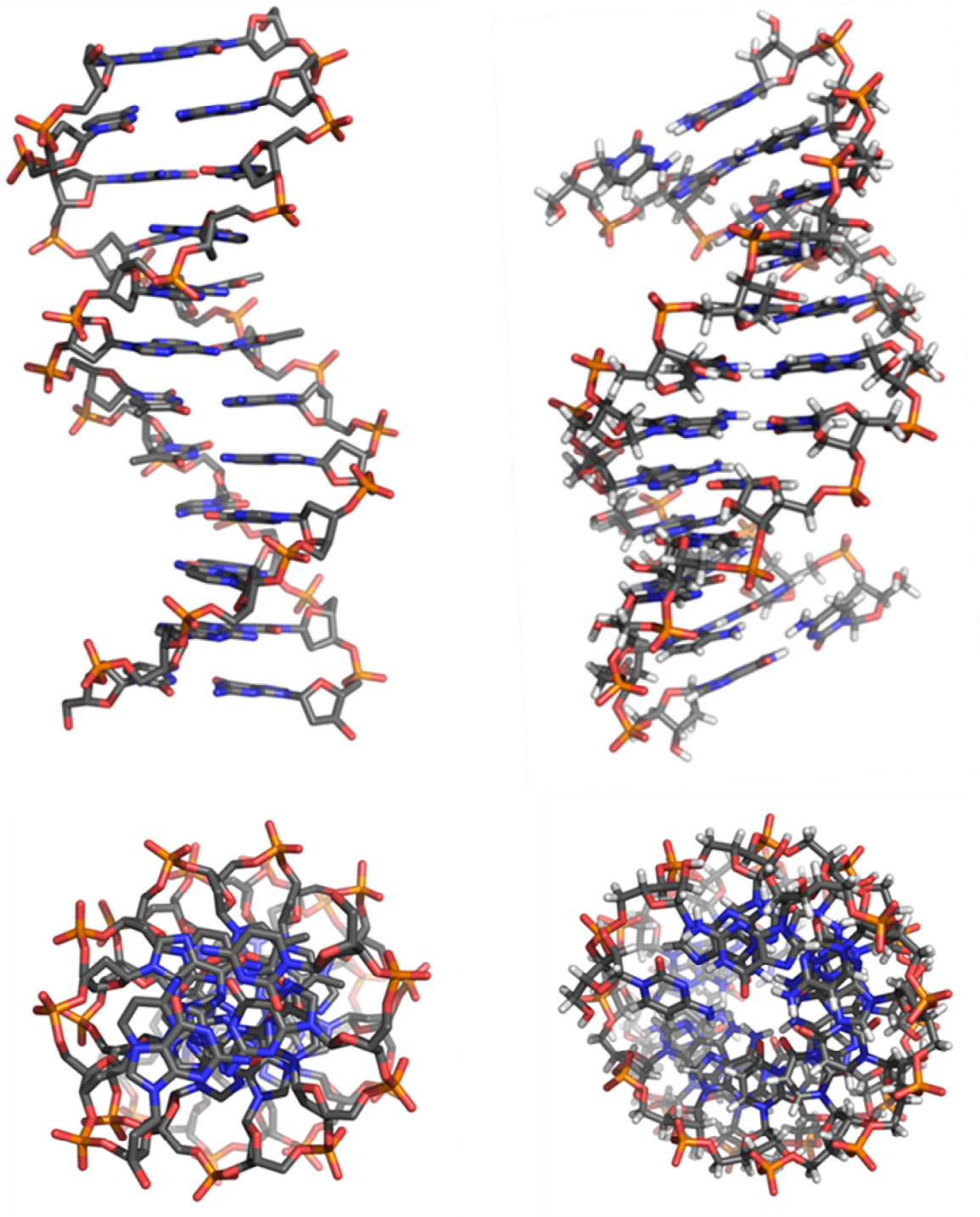
Figure 1.7. Typical helical structures for a) DNA and b) RNA, taken from the RSCB protein data bank, PDB ID: a) 1BNA and b) 1AL5. View along the helical axis is shown below each structure.

1.3 Triplex Structures

1.3.1 Historical background

Triple stranded nucleic acid structures were first discovered in 1957 by Felsenfeld and Rich[10], who found that a poly r(A) and poly r(U) sequence in the presence of 0.01 M $MgCl_2$ formed a complex with a stoichiometry of 1:2. They correctly proposed that a second poly r(U) strand was binding to the adenine strand by Hoogsteen base pairing[11] in the major groove of the duplex (Figure 1.8) made possible by the presence of divalent cations which neutralised the charges of the phosphate groups.
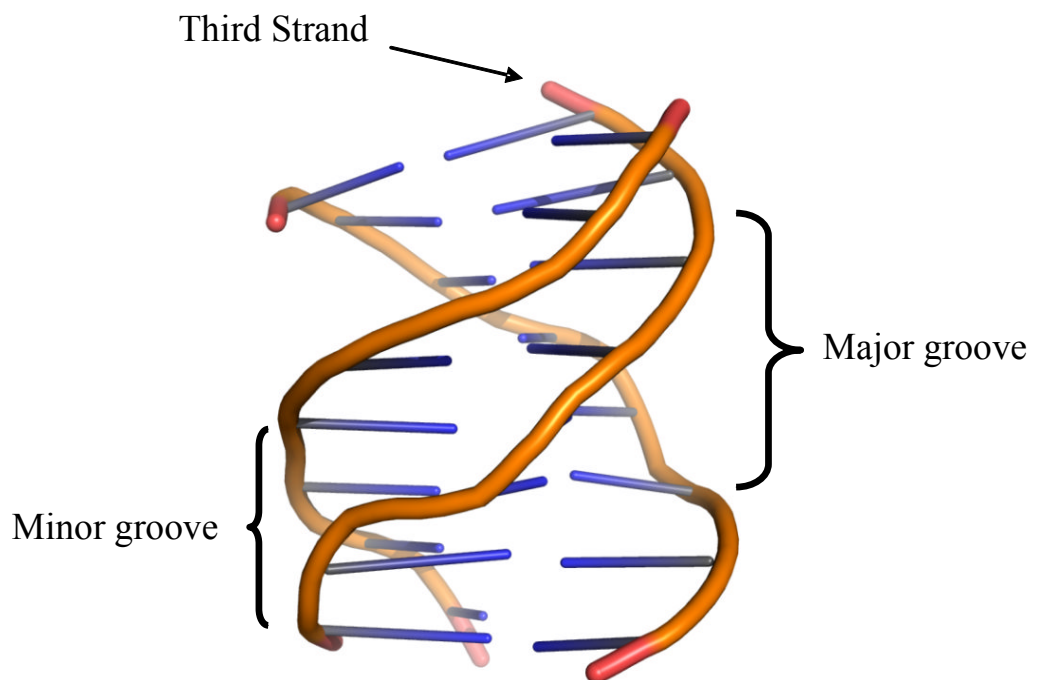
Figure 1.8. Solution structure of a DNA triplex showing binding of the third strand in the major groove. Image taken from the RSCB protein data bank, PDB ID 1D3X[12].

Additional triple-stranded RNA molecules were soon discovered; in 1963 Lipsett[12] showed that poly r(C) and poly r(G) could form a triplex, with a guanine third strand binding to an already formed GC duplex to give G:G-C base triplets (where ':' denotes Hoogsteen hydrogen bonds and '-' denotes Watson Crick base pairing). Lipsett also realised that at slightly acidic pH, the poly r(C) strand became protonated at the N3 (imino) positions, which enabled the formation of C^+:G-C base triplets[13].

Soon after, triple stranded DNA structures were demonstrated[14], as well as triplexes composed of both RNA and DNA polymers[15]. All of these studies involved homonucleotide polymers, however in 1968 it was realised by Morgan and Wells[16] that a stable triplex could be constructed from the mixed sequence poly r(UC):poly d(GA)-poly d(TC). This led to the realisation that triplex formation is sequence specific, and is restricted to polypurine-polypyrimidine sequences[9].

Structural studies of triplexes were first performed by Arnott in 1974 on the sequence poly d(T):poly d(A)-poly d(T) using x-ray diffraction data[17]. His studies showed that the triplex had an almost identical helical radius to the corresponding duplex (poly d(A)-poly d(T)). This led him to suggest that the triplex formed exhibited A-type geometry, which has a particularly deep major groove that could accommodate the third strand without disrupting the double helix. It has now been realised that the actual structure of a nucleic acid triplex depends on the sequence and type (i.e. DNA or RNA), but generally is somewhat intermediate between an A- and B-type helix[9].

Since the discovery and early studies of nucleic acid triplexes, it has been recognized that the exact structure of these triple stranded molecules can vary dramatically. Triplexes can be composed from DNA, RNA or a combination of the two, and can be formed from a single molecule (intramolecular) or from separate strands

(intermolecular). Two triplex motifs have also been described; namely the parallel and antiparallel motifs. These differ in the direction in which the third strand binds to the purine strand of the target duplex and are determined by the sequence composition of the third strand.

1.3.2 Parallel motif

Parallel or pyrimidine triplexes are formed when the third strand oligonucleotide is composed of pyrimidine residues. The third strand binds parallel (i.e. in the same 5′ to 3′ orientation) to the purine strand of the duplex, and parallel triplexes are characterised by the formation of T:A-T and C⁺:G-C triplets[18, 19] (Figure 1.9). Binding affinity in the parallel motif is dependent on the base composition of the third strand. TFOs which comprise only thymidine bases can form stable triplexes under physiological conditions, whereas TFOs containing cytosine bases form triplexes which are favoured at low pH, due to the requirement for cytosine protonation at the N3 position[20].

$$\textbf{T:A-T} \qquad\qquad \textbf{C}^+\textbf{:G-C}$$
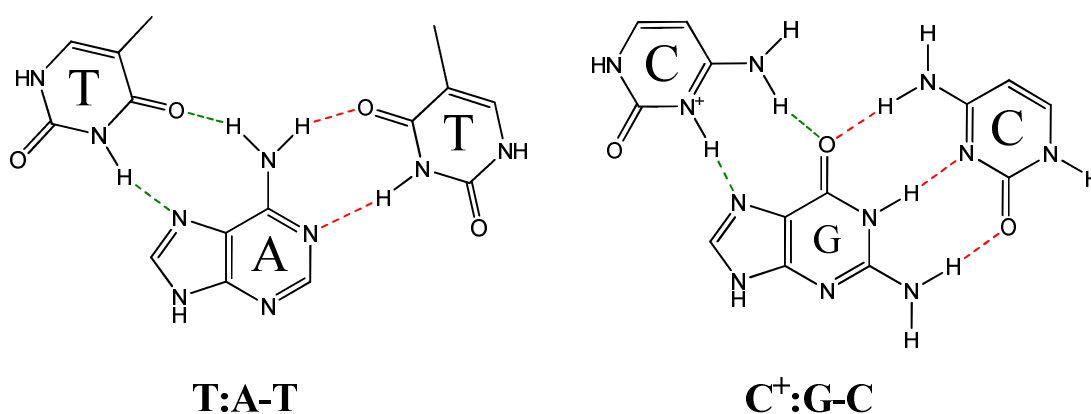
Figure 1.9. Triplet base pairing schemes for the parallel motif. Watson-Crick hydrogen bonds are shown in red, Hoogsteen hydrogen bonds are highlighted in green.

Parallel triplexes in which all three strands are composed of DNA oligonucleotides tend to retain a more B-DNA like structure, with the majority of sugars adopting the south pucker[21]. The inclination of the base pairs in a DNA triplex is small, which is similar to B-DNA. However, the helix is slightly unwound in comparison to that of an equivalent DNA duplex; this is necessary in order to accommodate the third strand in the major groove[9]. DNA triplexes are usually less stable than the underlying duplex, due to the repulsion of an extra negatively charged phosphate backbone, and the low pH requirement that is necessary for C^+:G-C triplets[22]. Recent results by Rusling *et al.*[22] have indicated that the stability of triplex DNA is dependent on the length and sequence of the underlying duplex, and that a duplex with enhanced stability will generate a more stable triplex.

RNA parallel triplexes are less stable than the corresponding DNA structures, but can be stabilised by the presence of divalent cations such as Mg^{2+}, due to electrostatic shielding of positive charges[23, 24]. Hybrid triplexes which contain both DNA and RNA strands have also been investigated, with UV absorption spectrophotometry and molecular modelling studies indicating that triplexes formed from a DNA duplex and an RNA third strand possess even greater stability than the corresponding all DNA structure[25]. The structure of these hybrid triplexes exists somewhere in between A- and B-form depending on the strand composition, usually containing a mixture of south and north sugar puckers[26].

1.3.3 Antiparallel motif

TFOs which contain only purine bases form antiparallel, or purine triplexes by binding antiparallel (in the opposite 3′ to 5′ direction) to the purine strand of the

target duplex. Although antiparallel triplexes are not characterised as definitively as parallel triplexes, it is known that binding of the third strand is by reverse Hoogsteen base pairing to form three possible base triplets, G:G-C, A:A-T and T:A-T[18, 19] (Figure 1.10). In contrast to parallel triplexes, antiparallel triplexes require no protonation and so exhibit pH independent binding[20].
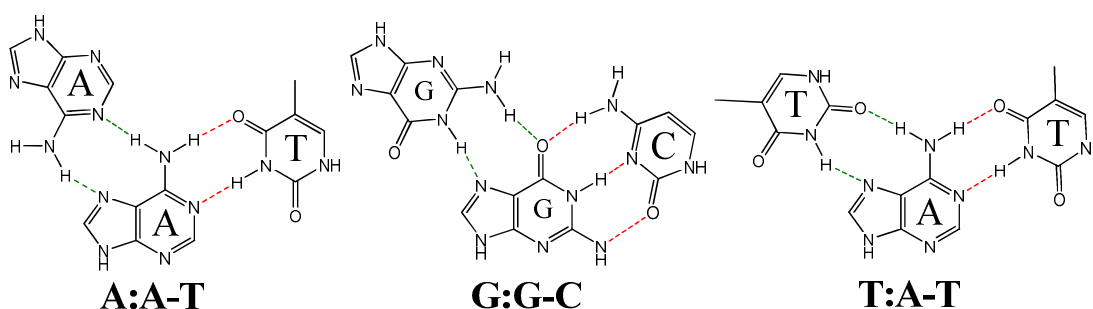
Figure 1.10. Triplet base pairing schemes for the antiparallel motif. Watson-Crick hydrogen bonds are shown in red, reverse Hoogsteen hydrogen bonds are highlighted green.

The structure of DNA antiparallel triplexes is more B-form like and the helical parameters and base pair inclination are generally similar to those of parallel triplexes[27, 28]. Very little is known about the structure of RNA antiparallel triplexes, however IR studies have indicated that they adopt an A-form like structure and contain predominantly north sugar puckers[9].

1.3.4 Intramolecular triplexes

Intramolecular triplexes can be formed from short mirror-repeat sequences with a single stranded loop, known as H-DNA or *H-DNA[29]. H-DNA involves half of the pyrimidine strand folding back on itself to form Hoogsteen hydrogen bonds with the

purine strand (Figure 1.11), whereas *H-DNA involves the purine strand folding back and interacting with the pyrimidine strand[9]. Two isomers of H-DNA are possible depending on whether the 3′ or 5′ half of the strand participates in triplex formation[30].
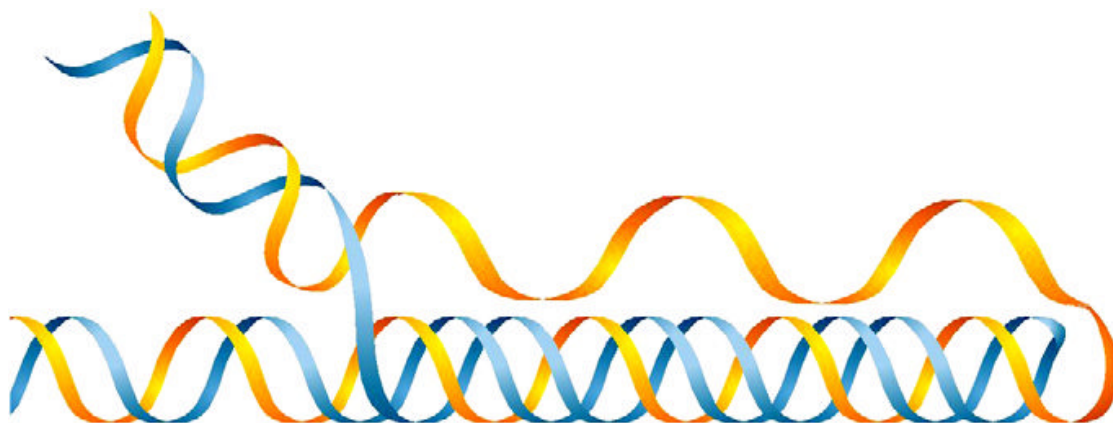
Figure 1.11. H-DNA[31]. The polypyrimidine strand (blue) folds back to form a triplex structure, and the polypurine strand (yellow) is left unpaired.

1.3.5 Intermolecular triplexes

Intermolecular triplexes are formed when a triplex forming oligonucleotide (TFO) binds to a target sequence on duplex DNA or RNA by Hoogsteen base pairing[18]. Binding of the third strand is to the purine sequence of the duplex, with the composition of the TFO determining the orientation in which it binds and under what conditions triplex formation is favoured.

1.3.6 DNA:RNA hybrid triplexes

The structure of triplexes which contain both DNA and RNA oligonucleotides exists somewhere in between A- and B-form like, depending on the strand composition,

with a mixture of south and north sugar puckers[26]. IR and chemical cleavage studies have shown a general trend of increasing A-form with increasing RNA content, and chemical cleavage and NMR investigations have indicated that the identity of the purine strand correlates with the helical conformation[32]. It follows that sequences with a DNA purine strand are likely to adopt more of a B-type helix, whereas triplexes containing an RNA purine strand will tend more towards an A-type helix.

In 1992, Roberts and Crothers examined the effects of backbone composition on triplex stability by mixing DNA-DNA, DNA-RNA and RNA-RNA 12mer duplexes with homopyrimidine DNA or RNA[33]. They found that a triplex only formed with a Hoogsteen DNA strand when it could bind to a purine DNA strand, and the most stable hybrid triplexes were those which consisted of a DNA duplex and an RNA third strand. Other UV absorption spectrophotometry and molecular modelling studies have since supported this finding, and have indicated that DNA-DNA:RNA triplexes possess even greater stability than their all DNA equivalents[25, 34-37]. It is thought that the conformational preference of ribose sugars to adopt the north pucker allows an RNA third strand to fit better into the major groove, causing less disruption to the Watson-Crick duplex[37].

Owing to this enhanced stability, hybrid triplexes have become a topic of interest over the past 20 years[35-38]. However, there are several problems associated with this type of triplex structure. Firstly, RNA synthesis is more longwinded than DNA synthesis, requiring longer coupling times due to the additional hydroxyl group at the 2′-position. 2′-protecting groups are therefore required, and must remain intact throughout the synthesis and then be removed specifically once synthesis is complete, without leading to chain migration or internucleotidic cleavage[38]. In addition, ribophosphoramidites have poorer coupling efficiencies than the equivalent

deoxyribophosphoramidites. As a result, more failure sequences are produced and it is only possible to synthesise relatively short RNA sequences, up to around 50 nucleotides in length[38, 39]. Finally, oligoribonucleotides are highly susceptible to degradation by ribonucleases, and tend to dissociate easily[38].

1.4 Modifications to improve triplex stability

Due to the enhanced stability of triplexes which comprise a DNA Watson-Crick duplex and an RNA third strand, modifications to the third strand of a DNA triplex have been developed to mimic RNA, to increase Hoogsteen bonding at physiological pH and also to increase the number of potential TFO target sites. There are three possible sites of modification on a nucleic acid, namely the base, the sugar ring and the phosphodiester backbone (Figure 1.12). Modifications to the base are somewhat limited due to the need to maintain Watson-Crick or Hoogsteen base pairs.
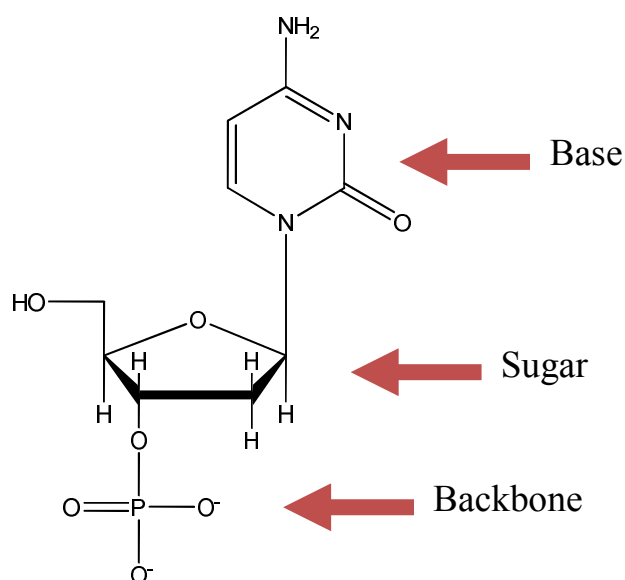
Figure 1.12. Three possible sites of modification on nucleic acids.

The acidic pH requirement of C^+:G-C base triplets is a major limitation of triplex formation at physiological conditions, and so a range of cytosine analogues have been developed in order to attempt to overcome this pH dependence. Substitution of 5-methylcytosine[40, 41] and 3-methyl-2-aminopyridine[42-44] (Figure 1.13) into a pyrimidine TFO has been achieved, both of which have a higher pK_a value than cytosine. In both cases an improved binding affinity of the TFO to a target duplex was observed as well as a reduction in C^+:G-C pH dependence. The enhanced stability can be attributed to stacking of the additional methyl groups within the major groove[41, 45] along with a larger degree of protonation at the N3 position as a result of the increased pK_a values. Other cytosine modifications have been developed which neutralise the positive charge at the N3 position[46-48], however these usually need to be used in combination with other stabilising analogues.
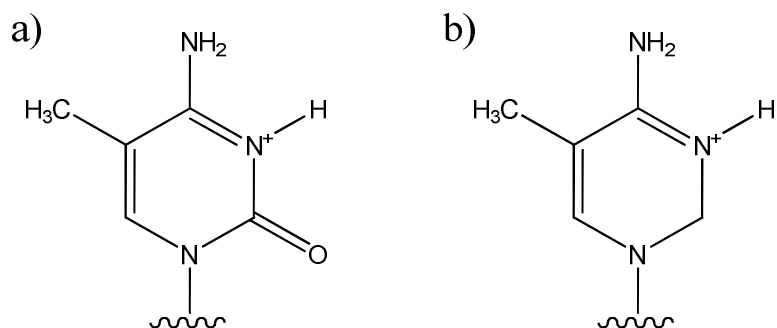
Figure 1.13. a) 5-methylcytosine and b) 3-methyl-2-aminopyridine residues.

The C^+:G-C triplet is more stable than the T:A-T triplet, due to screening of the charge on the phosphate groups by the protonated cytosine and favourable interactions with the stacked π-system[49]. A number of base modifications are aimed at improving the stability of the T:A-T base triplet, including pyrido[2,3-*d*]pyrimidine which increases the aromatic surface area[50] and 5'-propargylamino-

dU[19] which improves binding affinity by interacting with the adjacent 5′-phosphate on the third strand.

Base modifications to the purine strand of the duplex have also been investigated. Substitution of guanine by 6-thioguanine[51] and adenine by 7-deazaxanthine[52] were both found to help prevent the formation of unwanted intramolecular secondary structures.

A wide variety of modifications to the sugar ring have been explored, including a 2′-*O*-methyl analogue in which the 2′ proton on the sugar moiety is replaced with a methoxy group which has the function of altering the conformation of the ring from south to north, whilst providing resistance to degradation by ribonucleases[53] (Figure 1.14). Several bridged nucleic acid analogues have also been developed, such as the locked nucleic acid (LNA) in which a 2′-*O*,4′-*C* methylene bridge restricts or 'locks' the sugar into the C3′-endo conformation[54] (Figure 1.14). NMR studies by Sorensen *et al.* (2004) showed that a third strand consisting of alternating DNA and LNA residues had a greater binding affinity for duplex DNA than an equivalent DNA TFO, and produced a more stable triplex at physiological pH[55].
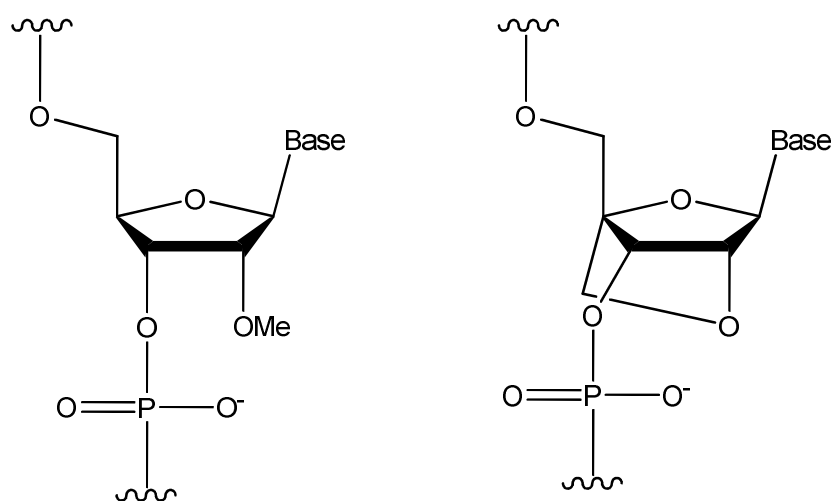
Figure 1.14. Chemical structures of a) 2′-O-methyl and b) LNA nucleotides.

Modifications to the phosphodiester group in the backbone of the third strand have also been widely explored. The phosphate backbone of an oligonucleotide is negatively charged, and so it is unfavourable to have three strands in close proximity. Backbone modifications which eliminate the negative charge on the phosphate have been investigated, such as Peptide Nucleic Acids (PNA), which have an achiral polyamide backbone consisting of N-(2-aminoethyl)glycine units, with the bases linked via a methylene carbonyl linker[54] (Figure 1.15). It has been reported that PNAs can bind to DNA duplexes by several different modes. They can form conventional Hoogsteen-type triplexes by binding in the duplex major groove[56], or triplex invasion complexes where the PNA displaces one of the PNA strands to give an internal PNA:DNA-PNA triplex[56, 57]. Triplexes containing uncharged morpholino units in the third strand (Figure 1.15) have also been probed and found to significantly promote pyrimidine motif triplex formation at neutral pH[58, 59].
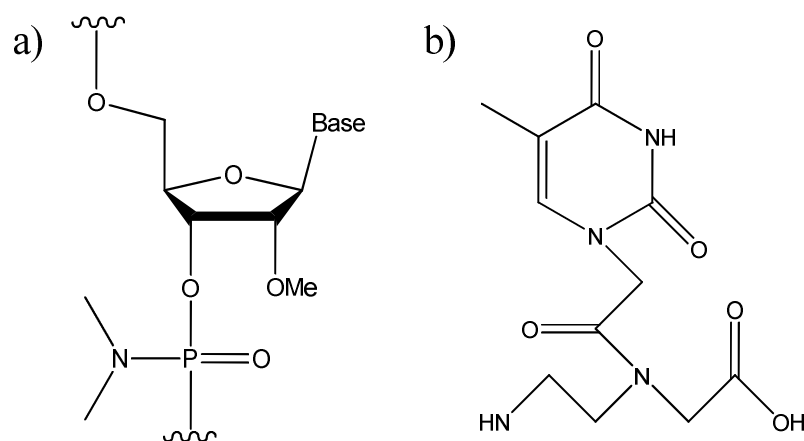
Figure 1.15. Chemical structures of a) morpholino and b) PNA subunits.

1.5 Biological significance of triplex motifs

1.5.1 Evidence for triplexes *in vivo*

There is considerable evidence for the existence of triplexes *in vivo*. For example, triplex structures have been detected in the nucleus of cells from insects, nematodes and mammals by immunofluorescence[60], and such structures have been found to be abundant in human cell nuclei[61]. This suggests that triplexes may have biological functions and until recently these were largely unknown. However, thanks to recent work on new methods for triplex detection, triple helices have now been implicated in a diverse range of biological processes such as transcription[62], RNA splicing[63, 64] and chromatin organisation[65].

1.5.2 *In vivo* applications of triplex DNA

Triplex forming oligonucleotides are promising tools for a variety of applications, due to their ability to recognise double-stranded DNA targets with high affinity and in a sequence-specific manner[18, 66]. Triple-helix target sites (i.e. stretches of homopurine-homopyrimidine tracts) are found in both prokaryotes and eukaryotes, and are often over-represented in eukaryotic sites[9].

One of the most promising potential applications of triplexes is in the control of gene expression, either by promoting or inhibiting gene transcription using a triplex forming oligonucleotide as a drug to bind to a duplex target within a gene. The use of TFOs is limited to the requirement of a homopurine-homopyrimidine target sequence and the stability of triplex structure formed, however as there is a need for agents which act on specific DNA sequences in order to recognise genes that are associated

with certain diseases there is considerable interest in the potential use of TFOs in gene regulation[67]. TFOs are ideal for this purpose as they are very specific DNA binders and are also very easy to synthesise.

1.5.2.1 Regulation of transcription: antigene therapy

In a living cell, genetic information is transferred from double helical DNA to messenger RNA (mRNA) during transcription, and then from mRNA to protein during translation[68]. So-called antisense technology blocks the translation step by using an oligonucleotide to target and bind to mRNA thus preventing protein synthesis. In contrast, an 'antigene' approach can cause transcription inhibition by employing TFOs to bind to duplex DNA, producing a triplex which prevents unwinding of the DNA duplex and so mRNA cannot bind, effectively stopping the flow of genetic information[68] (Figure 1.16).
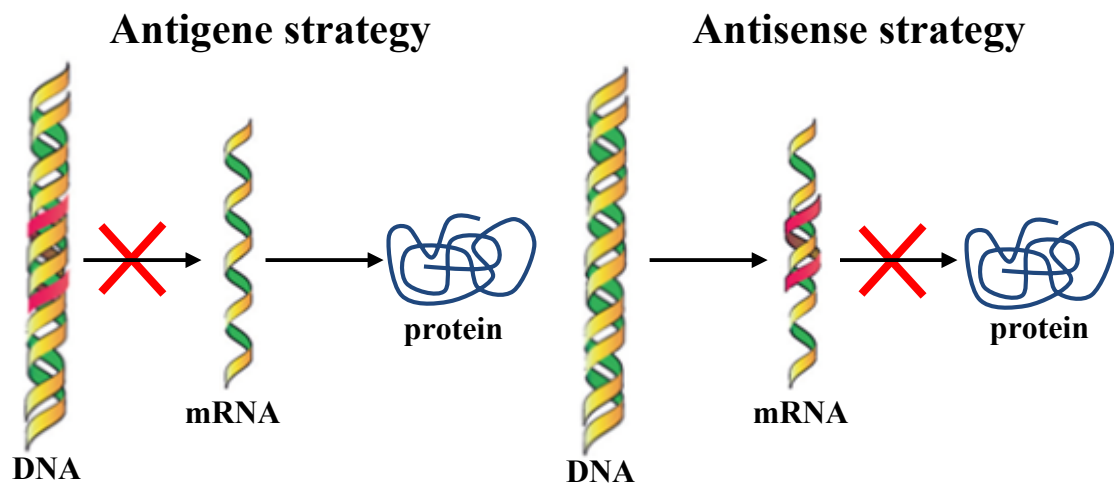
Figure 1.16. A schematic representation of antigene and antisense strategies for control of gene expression.

A major advantage of targeting the gene itself during transcription (i.e. antigene rather than antisense technology) is that there are only two copies (alleles) of a targeted gene compared to thousands of copies of messenger RNA[69]. Therefore targeting the gene itself can be a long term method for controlling gene expression, as it will prevent binding of mRNA for the residence time of the antigene oligonucleotide on its target sequence. This is in contrast to the antisense approach, which is a more short term method and requires replenishment of the antisense oligonucleotide[18]. Presently, most drugs are targeted at the protein level, so genetic information is still intact to produce more protein once the drug is withdrawn, therefore symptoms can return. If the drug is targeted at the DNA level and if inhibition can be held for the lifetime of the cell then the targeted gene will in principle, be removed in the next generation of cells. This strategy has potential for use against a multitude of viruses[18].

Other advantages of antigene technology include facile synthesis of the reagents and a wide availability of chemical modifications (to the bases, sugar-phosphate backbone and/or the 5′ and 3′ ends of oligonucleotides) which can improve cellular uptake, target binding, specificity and stability[18]. This will be discussed in more detail in Section 1.6.

Many examples of successful transcription inhibition have been reported for plasmid harboured genes[70, 71] and in several endogenous genes[67, 72-75]. In 2003, Carbone *et al.* designed a TFO with a high binding affinity for a specific homopurine-homopyrimidine sequence within the human Ets2 gene. This gene is known to have a role in cancer cell development, and binding of the TFO was found to significantly reduce Ets2 promoter activity and expression of the gene in prostate cancer cells[72]. A more recent study by Hewett *et al.* investigated using triple-helical DNA to inhibit

the tumour-associated tie-1 receptor gene[73]. The tie-1 promoter region was found to contain homopurine-homopyrimidine stretches of DNA, and binding of a TFO to these regions inhibited tie-1 activity *in vitro* and in endothelial cells by up to 75%[73].

It is also possible for triplexes to act as promoters of gene transcription, in order to increase expression of genes that are transcribed at low levels. This can be used for treating some genetic diseases such as the β-globin disorder, sickle-cell anaemia[76, 77]. In 2004 Song *et al.* designed a psoralen-TFO adapter which was able to deliver an artificial enhancer to a disabled gene and activate transcription of that gene by up to four fold[78], and in 2005 Ghosh *et al.* demonstrated that a hairpin-TFO was able to activate expression of two weakly expressed genes in *Saccharomyces cerevisiae*, a strain of yeast[79].

1.5.2.2 Modulation of DNA replication

Several reports have indicated that TFOs are capable of inhibiting DNA replication[80-82]. For example, Pesce *et al.*(2005) demonstrated that a PNA oligonucleotide was able to inhibit HIV-1 production in lymphocytes and macrophages by preventing replication of the HIV gene[80]. More recently, it has also been shown that in some circumstances, triplexes can enhance DNA replication[83]. DNA polymerases are able to bind three DNA strands in their catalytic centre, enabling TFO primers bound to a duplex in a triple helix manner to be elongated[83].

1.5.2.3 Triplexes as biotechnological tools

TFOs can be employed in the targeting of therapeutic agents to specific sites of DNA in order to induce their effects in a sequence-specific manner. The first example of

this was reported by Matteucci *et al.* in 1997, who reported that the Topoisomerase I (Topo I) enzyme could be inhibited at specific DNA sites with the use of a TFO[84]. Topo I is present in all mammalian cells and its function is to relax the superhelical twist in DNA which is generated during transcription and replication. This process is inhibited by the antitumour agent camptothecin (CPT) binding to the DNA enzyme complex and preventing re-ligation. Topo I poisons display weak sequence specificity, with very few recognition elements for CPT[85], and so by tethering camptothecin or one of its analogues to a TFO, the binding of the inhibitor to the enzyme complex can be made to be sequence-specific, enabling the inhibition of Topo I to be controlled[84]. Other inhibitors have since been used for this approach, such as rebeccamycin[86] and indolocarbazole[87], and the same principle has now been applied to topoisomerase II inhibitors[88].

Another possible application of the triplex approach is the recognition and purification of nucleic acids. In a mixture of duplexes, homopurine-homopyrimidine sequences that are able to form a triple helix can be identified and separated by introducing a TFO attached to magnetic beads or an affinity column[89]. Schluep and Cooney published a successful method for affinity purification of plasmids by triplex interaction (Figure 1.17), which included immobilising a pyrimidine TFO on to a large pore chromatography support[90]. Plasmid DNA containing a suitable target for the TFO was successfully purified by incubating with the support, showing no RNA or DNA contamination in HPLC analysis[90]. One advantage of using a pyrimidine TFO for this purpose is the pH dependence of parallel triplexes, which can allow the purification to be tuned[67].
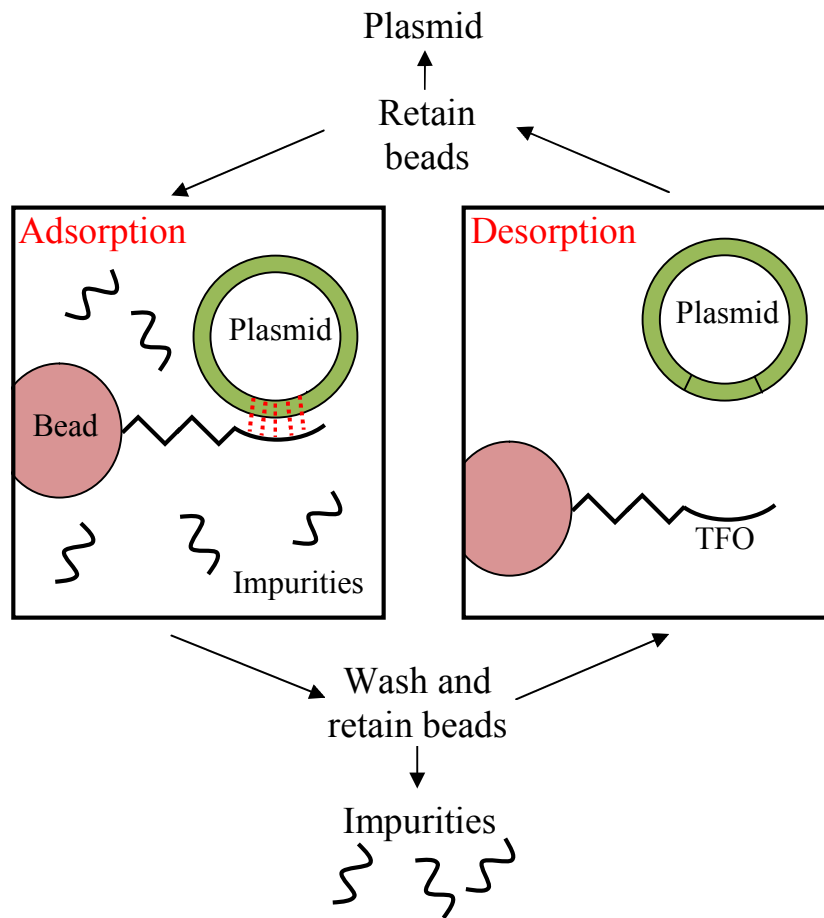
Figure 1.17. Method for purification of plasmid DNA by triplex affinity capture.

1.6 Proposed investigation

As described in Section 1.5.2 there are several promising biological and therapeutic applications of triple helical nucleic acid structures. However, several problems must be overcome in order to maximise the potential use of triplexes in for example, gene therapy. Such problems include the pH limitation for the formation of cytosine-containing parallel triplexes and the synthesis difficulties associated with RNA third strands despite the enhanced stability of a triplex comprising a DNA duplex bound to an RNA third strand. It would therefore be of great advantage to devise an all-DNA triplex in which the third strand is RNA-like, so as to overcome the problems associated with RNA oligonucleotides, but to improve the stability of the structure.

1.6.1 3′-S-phosphorothiolate modification

Several studies have previously incorporated 3′-S-phosphorothiolate linkages into a DNA duplex, which involves replacing the 3′-oxygen in the phosphodiester backbone with a sulfur atom in one or more of the nucleotides[91-95]. The effect of this is to cause the structure to adopt a local RNA-like conformation at the site of modification. Replacement of the 3′-bridging oxygen with sulfur results in a conformational shift from a DNA-like south, to an RNA-like north conformation of the deoxyribose ring to which it is attached and the sugar ring 3′ (n+1) to it[92, 95] (Figure 1.18).
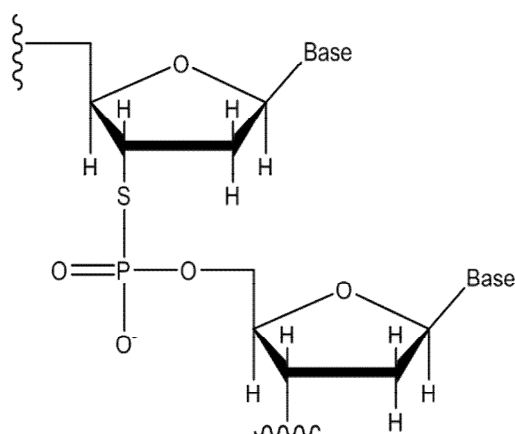
Figure 1.18. A 3′-S-phosphorothiolate linkage.

The conformational shift is due to *gauche* effects within the deoxyribose ring, particularly that between the 3′-substituent and the ring oxygen which is optimised in the south conformation[95]. Sulfur is sterically larger and less electronegative than oxygen, and reduces the influence of these *gauche* effects. This forces the conformational equilibrium to move towards the north conformer as governed by the anomeric effect[92, 95, 96].

Beevers *et al.* reported that a single 3′-*S*-phosphorothiolate link incorporated into the DNA strand of a DNA-RNA duplex causes significant localised conformational changes. NMR experiments on both unmodified and modified duplexes were performed, and J-couplings were measured in order to calculate the approximate mole fraction of south conformer for each residue. The conformation of the modified residue was found to have shifted from south to almost completely north, and the sugar (3′) to the modification showed a similar shift, although to a lesser extent[97]. An increase in thermal stability of 2.5 °C was also observed for the modified duplex, compared to the unmodified equivalent[97].

More recent studies have investigated the effects of several 3′-*S*-phosphorothiolate linkages incorporated into one DNA strand. Jayakumar *et al.* investigated the effects of two consecutive modifications within a deoxythymidine trinucleotide and found that sequential 3′-*S*-phosphorothiolate links result in a loss of conformational change of the 3′ ring on the (n+1) residue[92]. This is thought to be due to interactions between the bases of sequentially modified subunits[92].

Bentley *et al.* have recently reported that systems containing alternate 3′-*S*-phosphorothiolate linkages have increased stability compared to those containing consecutive 3′-*S*-phosphorothiolate linkages[92, 96]. This is because alternate modifications cause more sugar moieties to undergo a south-to-north conformational shift[96], unlike sequential modifications, in which the 3′ conformational flip from south to north is lost.

3′-*S*-phosphorothiolate linkages have been incorporated into an i-motif structure[93] (Brazier *et al.*, 2006). However, to this date, there have been no studies of triplex structures containing 3′-*S*-phosphorothiolate modifications.

1.6.2 Project aims

The aim of this project will be to synthesise and compare two equivalent DNA triplexes – one unmodified and the other containing modifications in the third strand which will 'mimic' RNA (i.e. A-form, north sugar conformations). $3'$-S-phosphorothiolate linkages will be used as these are known to provide resistance to nuclease degradation and have not previously been incorporated into triplex structures. The $3'$-S-phosphorothiolate analogue is advantageous as it is easily synthesised[91] and is stable in a physiological environment[98], making it an ideal candidate to be exploited for possible biological or biotechnological purposes. It is anticipated that the modified third strand will exhibit an increase in binding affinity for the target hairpin, producing a more stable triplex than the non-modified equivalent.

1.6.3 Triplex to be analysed

The DNA system under investigation will form a bimolecular triplex, and will be composed of a hairpin duplex and a single stranded TFO. This was preferred over an intramolecular system as the potential biological and therapeutic uses of DNA triplexes would involve a TFO recognising and binding to a duplex target.

The hairpin must be a potential target for a TFO, and so must consist of a homopurine-homopyrimidine sequence. A base sequence that is part of a known triplex-forming oligonucleotide[25] was selected and is shown in Scheme 1.1. The two strands of the hairpin are connected by a commercially available triethylene glycol loop known as spacer-9 (Figure 1.19). Spacer-9 is used in preference to a nucleic acid based loop in order to decrease the spectral complexity as well as to eliminate the

possibility of nucleotide loop residues becoming involved in base pairing which may affect the structure of the triplex[32]. A deoxyuridine residue was also included to break up the symmetry of the duplex. Uridine is equivalent to thymidine in that it will bind to adenosine, but it does not contain the methyl group so will be more easily distinguishable from other thymidine nucleotides in the NMR spectra. G-C base pairs were included at both ends of the hairpin, in order to minimise 'fraying' - where the base pairs at the ends of the molecule can open due to the intermittent breaking of hydrogen bonds[99]. G-C base pairs, which contain three hydrogen bonds help to minimise fraying as they are more stable than A-T base pairs, which have only two hydrogen bonds holding them together.

5′ - G A G G A A A G A A G G – spacer-9 – C C T T C dU T T C C T C - 3′

Scheme 1.1. The sequence of the target hairpin under investigation. Spacer-9 represents a triethylene glycol linker and dU denotes a deoxyuridine residue.
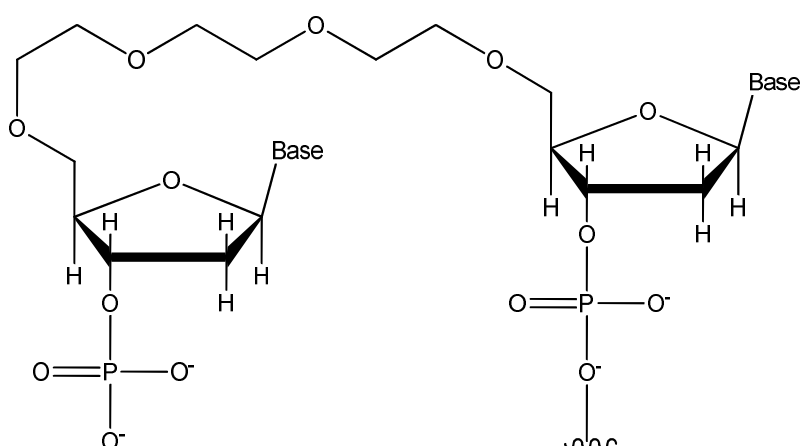
Figure 1.19. Structure of the triethylene glycol linker attached to two nucleotides.

The base sequence of the triplex forming oligonucleotides was selected so that they would bind to the hairpin in a parallel orientation as shown in Scheme 1.2. Several factors were taken into consideration when choosing the sites at which the modifications were to be introduced. Although 3′-S-phosphorothiolate linkers can be incorporated into any nucleobase[91, 100], the thymidine residue is easiest to modify[101] therefore it was decided to include the modifications between two thymidine nucleotides. It was also taken into consideration that the stabilising effect is maximised when 3′-S-phosphorothiolate modifications are placed in non-adjacent positions, as described in Section 1.6.1.

Natural: 5′ - C T C C T T T C T T C C - 3′

Modified: 5′ - C T C C T T **s** T C T **s** T C C - 3′

Scheme 1.2. The sequence of the natural and modified triplex forming oligonucleotides. T s T indicates a 3′-S-phosphorothiolate link.

1.7 Analysis techniques

1.7.1 NMR spectroscopy

Nuclear Magnetic Resonance (NMR) spectroscopy is an ideal technique for use in studying the structure of nucleic acids. Although structures can be obtained by x-ray crystallography, there are several drawbacks to this method such as difficulty in crystallising the sample[102]. Whilst x-ray crystallography provides the most precise atomic detail, information about the dynamics of the molecule is limited[103]. Since nucleic acids carry out their biological role under aqueous conditions and are

dynamic in solution, solution state NMR is the preferred technique for nucleic acid investigations.

The overall structure of a nucleic acid system can be determined by NMR spectroscopic methods; a variety of experiments can provide information such as chemical shifts and coupling constants which can subsequently be used to determine the sugar puckers and glycosidic bond orientation of individual nucleotides. Proton distances can also be extracted from nuclear Overhauser enhancements (nOes) which may be used as experimental constraints for generating high resolution structures. This is described in more detail in Chapter 3.

1.7.2 UV melting studies

Ultraviolet (UV) spectrophotometry is a useful tool for measuring the stability of a nucleic acid structure. The unfolding of a DNA secondary structure can be monitored by measuring the change in the UV absorbance at 260 nm with increasing temperature. Upon heating, π-π interactions between stacked base pairs become disrupted and the structure unfolds, causing an increase in UV absorbance known as hyperchromism[104, 105]. As the structure becomes more disordered, the absorbance of the oligonucleotide approaches that of the constituent nucleotides, therefore the degree of hyperchromicity is a qualitative measure of base pairing and stacking within a secondary nucleic acid structure[105].

For duplex DNA, a plot of the absorbance at 260 nm versus temperature produces a monophasic, sigmoidal curve which is similar in appearance to a phase transition, and so is termed a 'melting' curve[105] (Figure 1.20). The midpoint of the curve is defined as the melting temperature, T_m and is the point at which half of the DNA is in

its ordered structure and half is in its disordered (single stranded) state. T_m values measure the thermal stability of the secondary structures formed by DNA and can be determined by plotting the first derivative of the melting profile.
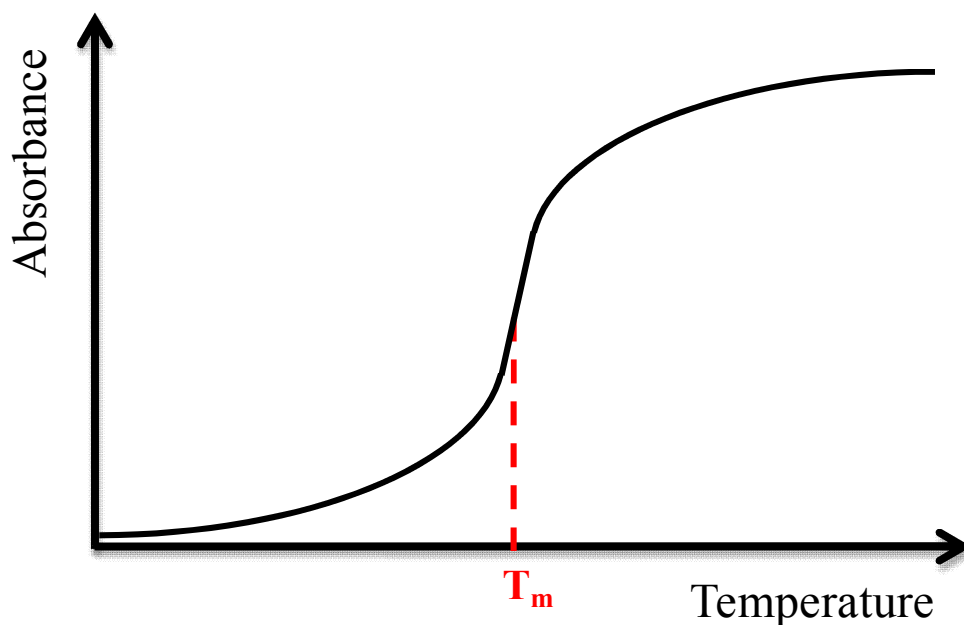
Figure 1.20. Schematic of a monophasic melting curve.

1.8 Plan of Investigation

This chapter has attempted to encompass all of the relevant background information to the proposed study. The following two chapters will describe the NMR analysis and structural investigation of the target hairpin, respectively. This will provide the basis for comparison to both the native and modified triplex NMR assignments (Chapter 4) and structural analysis (Chapter 5). The investigation will then be completed with the UV thermal melting analysis of the target hairpin and the natural and modified triplexes, which is detailed in Chapter 6.

Chapter 2

NMR Assignments of the Target Hairpin

Chapter 2.   NMR Assignments of the Target Hairpin

2.1 Introduction

As already discussed, the purpose of these studies is to establish the structural and thermodynamic consequences of 3′-*S*-phosphorothiolate modifications in the third strand of a parallel triplex. This will be achieved by analysing the NMR spectrum of the triplex, with and without the chemical modifications. In advance of studies of the two triplexes, it was necessary to analyse the spectrum for the hairpin target, shown in Scheme 2.1. Work towards the complete ^{1}H resonance assignment and qualitative structural assessment is described here.

Individual protons were assigned using a series of two dimensional NOESY, TOCSY and DQF-COSY spectra. The hairpin was analysed in D_2O solution to assign the non-exchangeable protons, and in a 9:1 H_2O:D_2O mix to enable assignment of the exchangeable protons. NOe volumes were measured from NOESY spectra recorded with a range of mixing times, from which inter-proton distances could be calculated[106] for use in subsequent structure generation for the hairpin (Chapter 3).

5'- G_1 A_2 G_3 G_4 A_5 A_6 A_7 G_8 A_9 A_{10} G_{11} G_{12} ⌐

spacer-9

3'- C_{24} T_{23} C_{22} C_{21} T_{20} T_{19} dU_{18} C_{17} T_{16} T_{15} C_{14} C_{13} ⌐
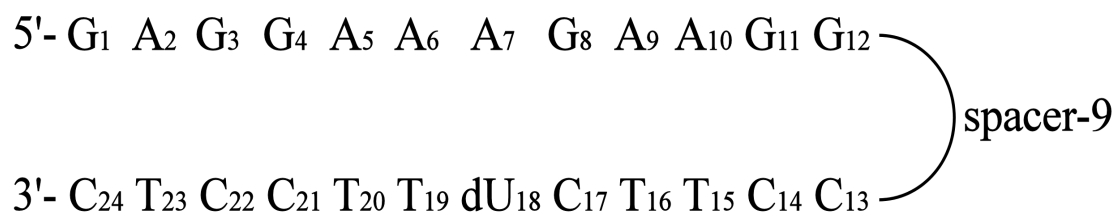
Scheme 2.1. The sequence and numbering of the hairpin under investigation.

2.2 Assignment of non-exchangeable proton resonances

Due to the polymeric nature of DNA, many of the assignable hairpin protons have very similar chemical shifts, resulting in significant crowding and overlap in several regions of the one dimensional spectrum. Although the 1D spectrum displays clearly defined regions of the different classes of protons (Figure 2.1), individual proton resonances cannot be distinguished, requiring the use of two dimensional techniques for a complete resonance assignment. The labelling of the non-exchangeable protons is shown in Figure 2.2.
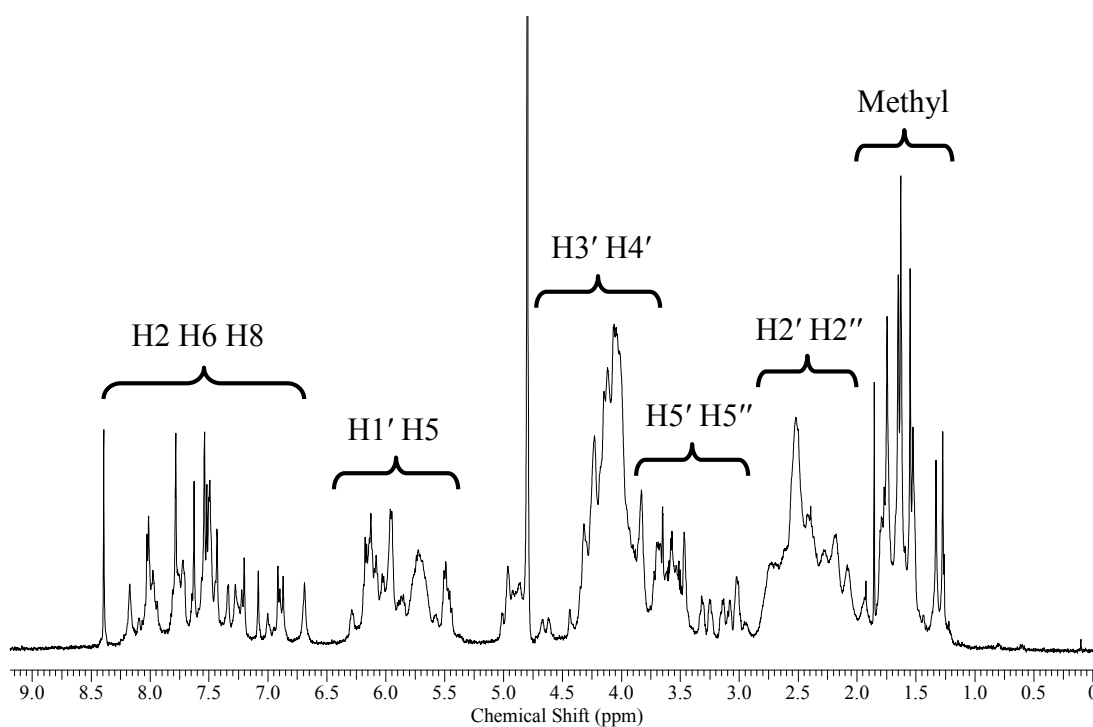
Figure 2.1. One dimensional ^{1}H NMR spectrum for the hairpin with proton chemical shift regions labelled.
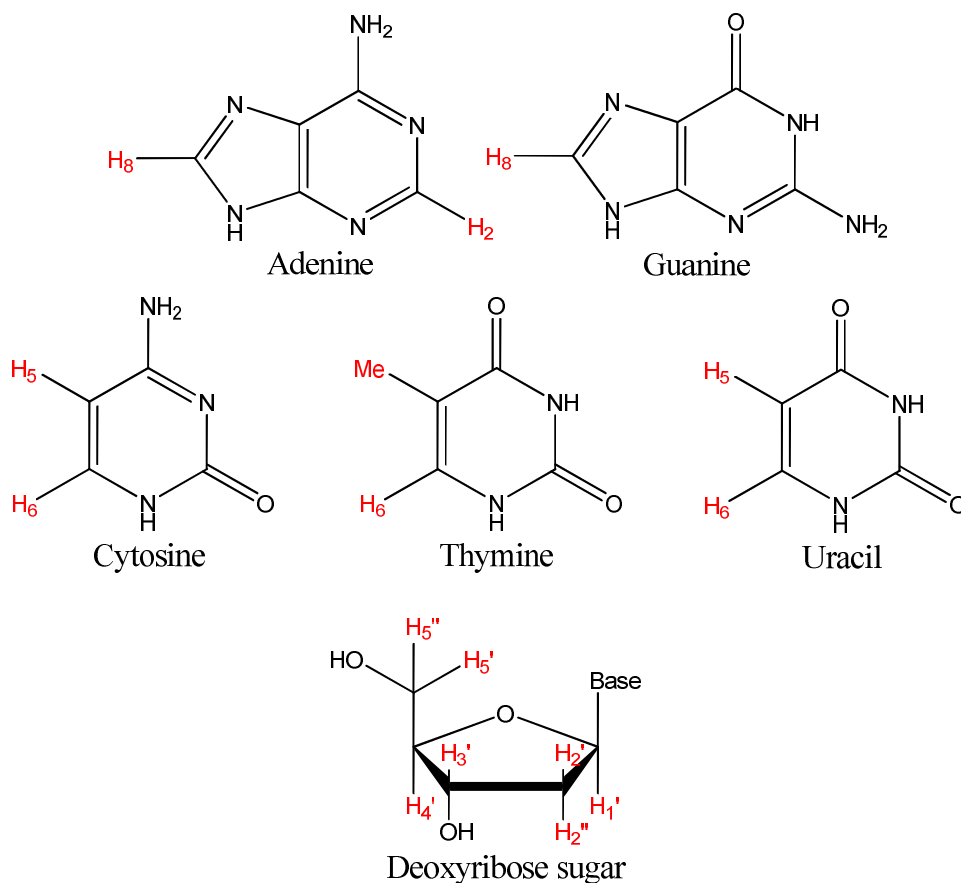
Figure 2.2. The structure of the deoxyribose sugar and heterocyclic base units. Numbering for the non-exchangeable protons is indicated in red.

A combination of DQF-COSY, TOCSY and NOESY experiments were used to assign the majority of non-exchangeable protons[107, 108]. The DQF-COSY spectrum allows couplings over two or three bonds to be observed, whilst the TOCSY identifies all through-bond couplings in an uninterrupted spin system. Figure 2.3 shows some of the expected TOCSY connections for a single cytosine nucleoside.
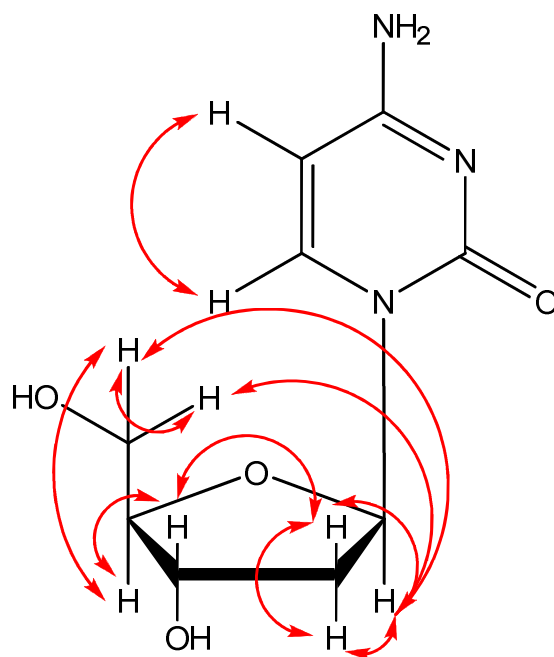
Figure 2.3. Expected TOCSY connections for a cytosine nucleoside.

The NOESY spectrum reveals through-space connections between protons that are separated by a distance of less than 6 Å[109], therefore it enables connections to be made between base and sugar protons which are not close in terms of bonds but which are close in space. A NOESY spectrum is imperative for the complete assignment of an oligonucleotide, as it exhibits inter-residue couplings which allows sequential assignments of residues to be made. As with the 1D spectrum, the NOESY spectrum can be divided into regions in which specific proton connectivities occur (Figure 2.4).
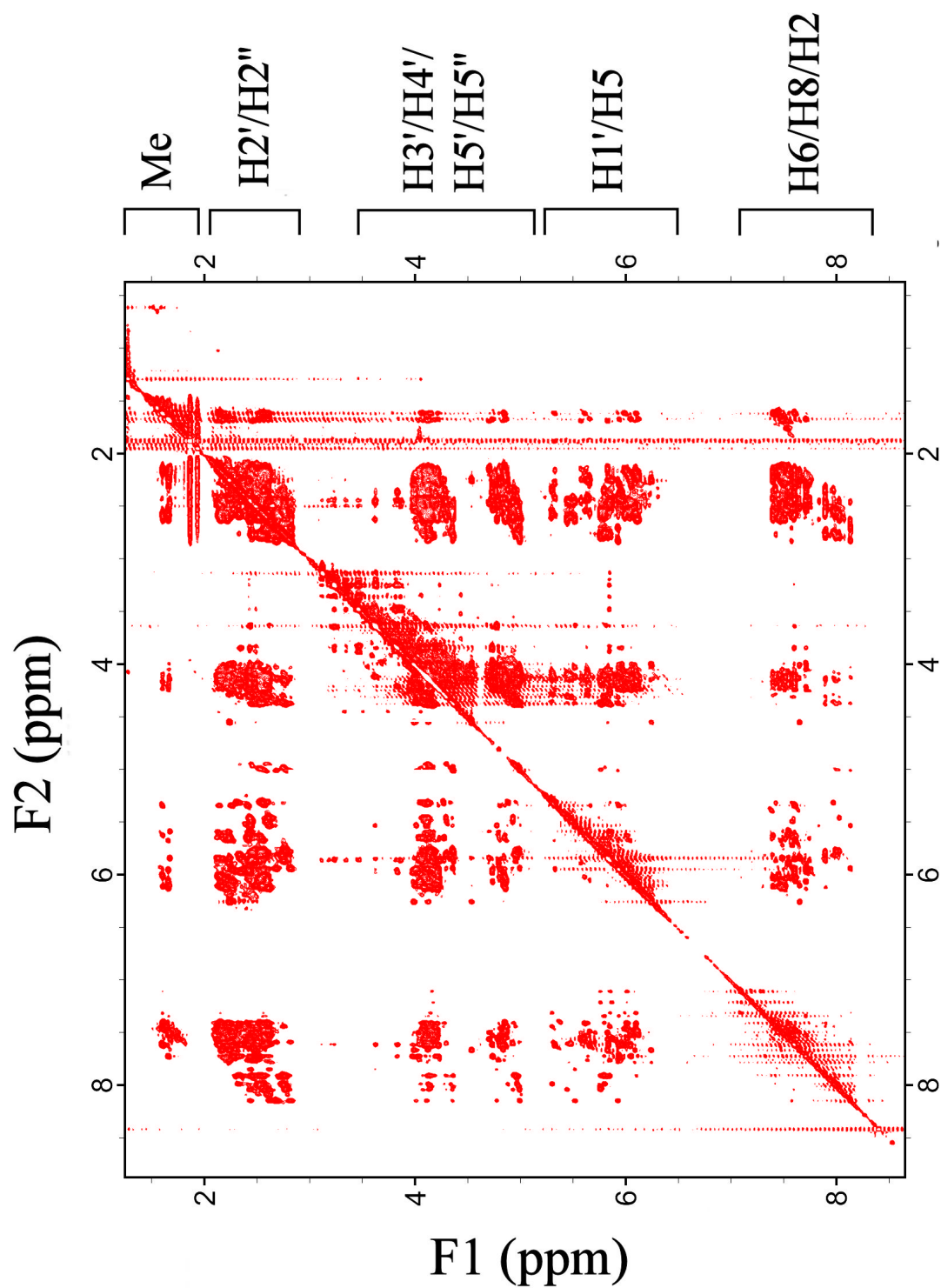
Figure 2.4. 500 MHz NOESY NMR spectrum of the hairpin in 20mM phosphate buffered D$_2$O solution at 293 K. The specific proton regions are labelled.

A useful starting point in assigning the hairpin spectrum is to identify the thymidine (T) Me-H6 and cytidine (C)/deoxyuridine (dU) H5-H6 cross peaks, as these regions are relatively uncrowded. Intra-residue H5-H6 cross peaks can be observed in the TOCSY and NOESY spectra, and additional inter-residue dipolar couplings are observed in the NOESY spectrum. In order to assign these we must first assume a model for the structure. The most appropriate model for the hairpin was the B-type double helix that is typical for DNA duplexes[1]. If this model was correct then the connections shown in Figure 2.5 would be seen[110], allowing sequential connections to be made between T/C/dU methyl/H5 protons to the H6 proton of a neighbouring base in the sequence.
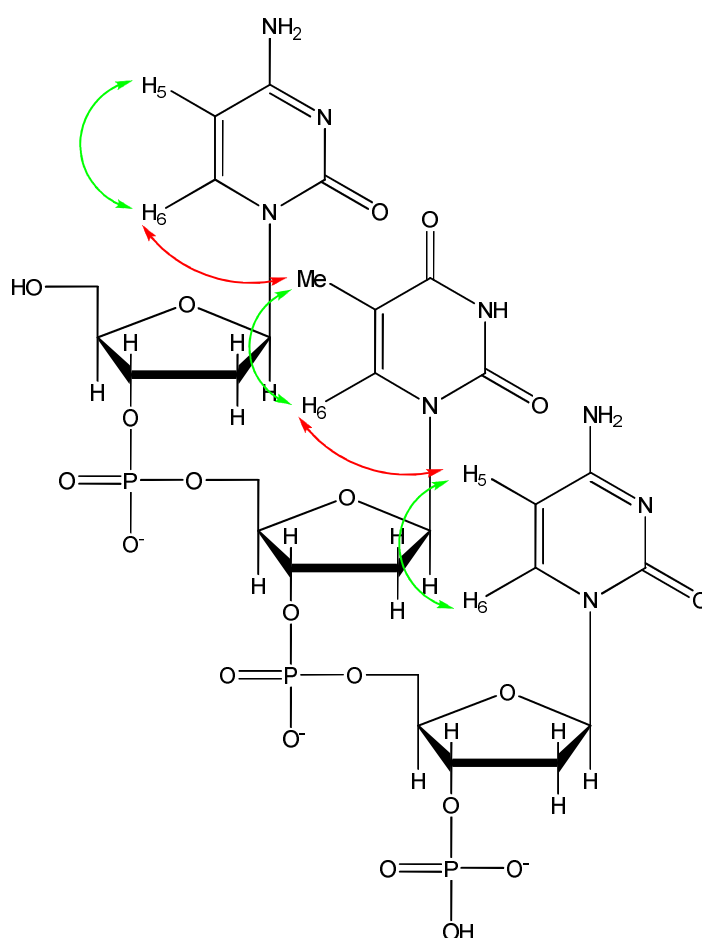
Figure 2.5. Expected TOCSY (green) and NOESY (red) connections between the Me, H5 and H6 protons of thymidine and cytidine nucleotides in a B-type DNA duplex.

As expected, there were five Me-H6 and seven H5-H6 TOCSY cross peaks, which correspond to the five thymidine and seven cytidine/deoxyuridine residues in the sequence (Figure 2.6).
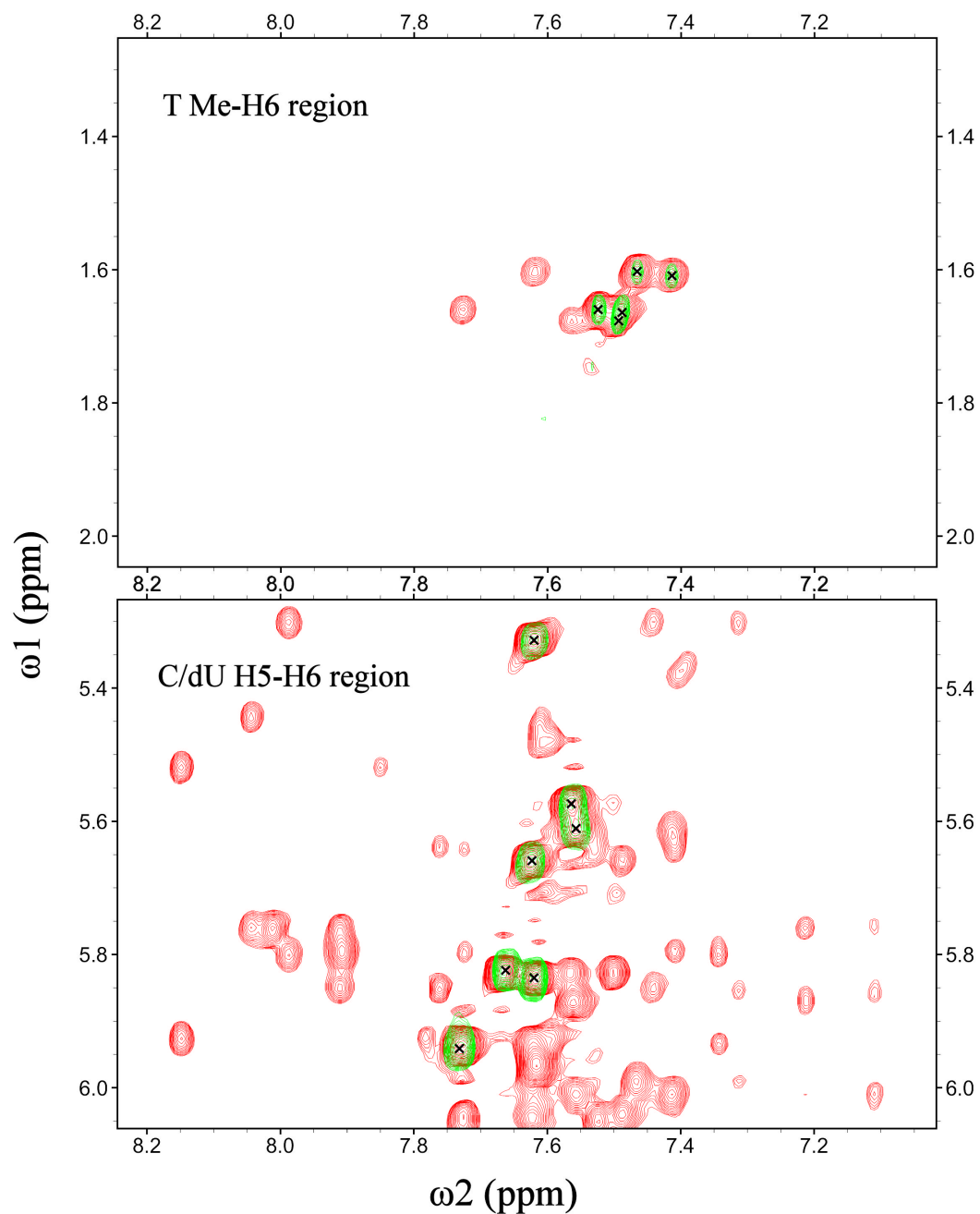
Figure 2.6. Expansions of the NOESY (red) and TOCSY (green) spectra showing the T Me-H6 and C/dU H5-H6 regions.

The sequences of residues can be identified from the inter-residue NOESY connections (Figure 2.5). Of the thymidine methyl groups, three exhibit connectivities to the adjacent 5′ cytidine H6 protons. Likewise, three of the cytidine H5 protons have an nOe correlation to the H6 of the thymidine Me (n-1) to it, allowing sequential assignments to be made (Figure 2.7).
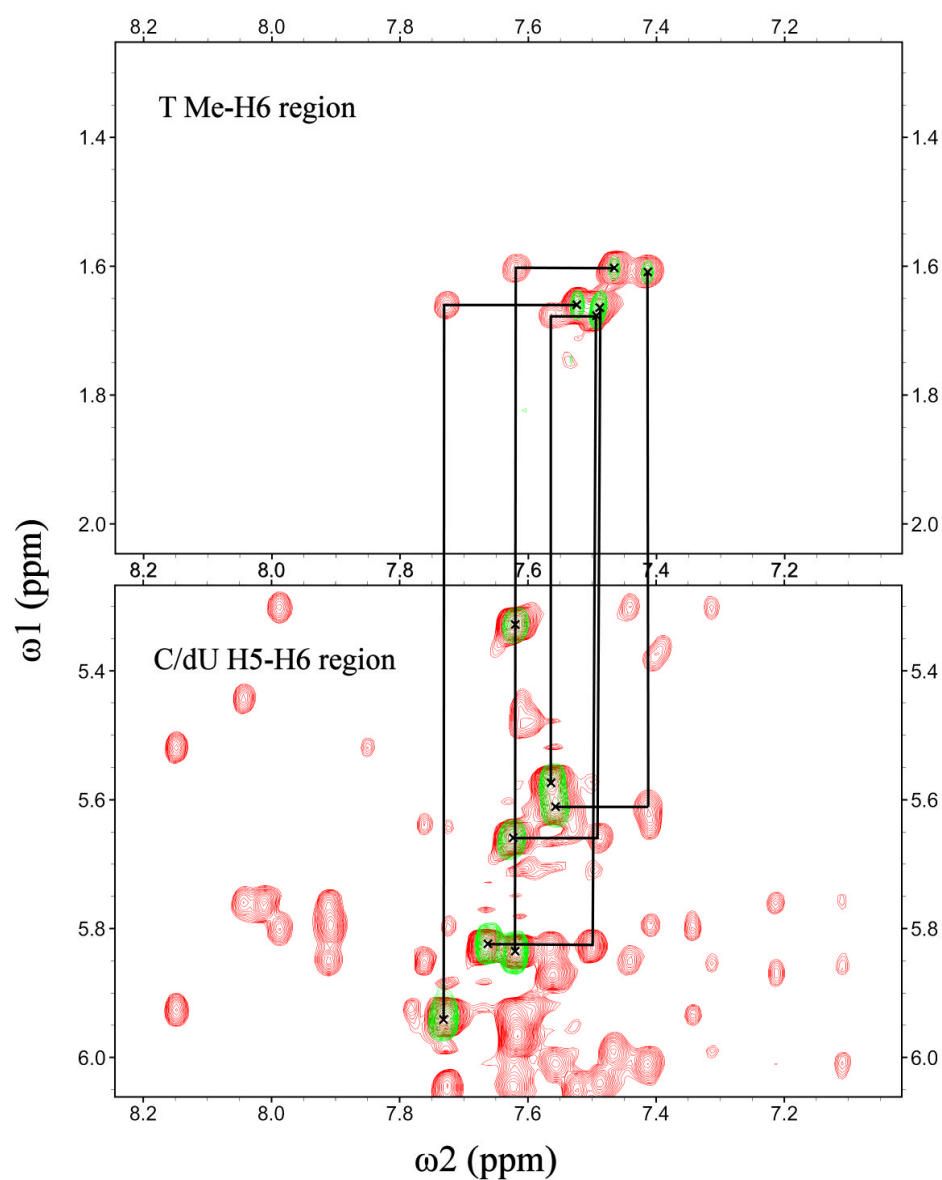
Figure 2.7. Expansions of the NOESY (red) and TOCSY (green) spectra showing inter-residue nOe connections.

Only one thymidine Me-H6 cross peak exhibits inter-residue nOe correlations to two different cytidine residues, and therefore must correspond to T23, which is the only T in the sequence which has both 3′ and 5′ cytidine neighbours. As T23 could be unambiguously assigned, so could C24 and C22. The H5 of C22 would display an nOe correlation to the H6 of C21, however this NOESY cross peak was absent from the spectrum. As there was another H5-H6 cross peak close to that of C22, it was assumed that this would correspond to C21, and that the inter-residue C22 H5-C21 H6 nOe is obscured underneath the much stronger C21 H5-H6 cross peak. A connection could then be made to T20 Me-H6, leaving three thymidine residues unassigned. Of these, only one has a cytidine residue 3′ to it (T16), so this could be assigned based on the T(n) Me-C(n+1) H6 NOESY cross peak. The two remaining thymidines, T15 and T19 should show connections to T16 and T20 respectively, however no nOe correlations were observed in the NOESY spectrum. It was assumed that these nOes were concealed underneath nearby intra-residue cross peaks, and so based on this assumption, the T15 and T19 Me-H6 interactions could be assigned (Figure 2.8). The remaining C/dU H5-H6 cross peaks were assigned via nOe connections to neighbouring thymidine residues (Figure 2.9).
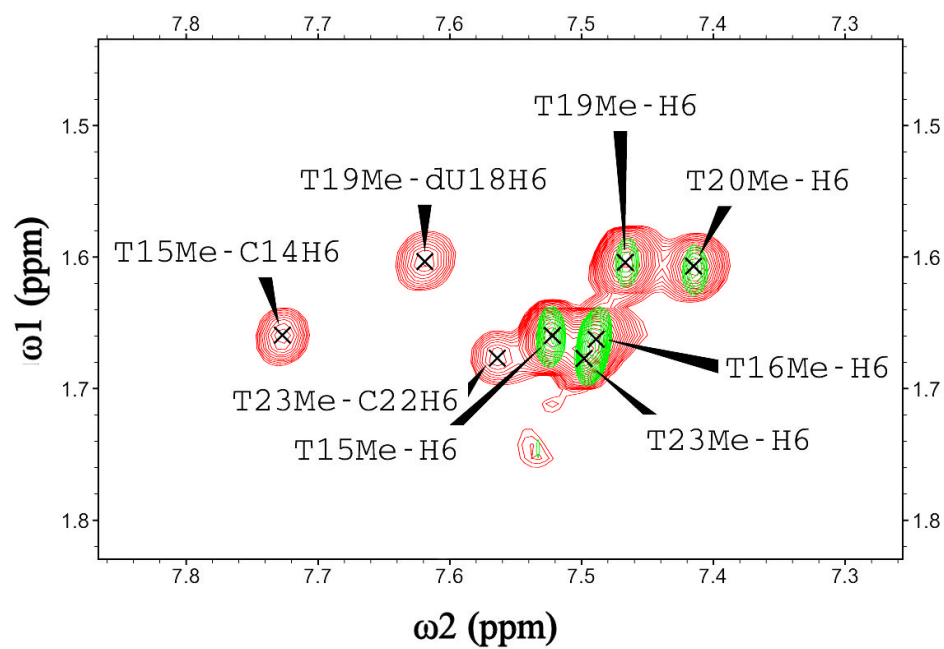
Figure 2.8. Expansions of the TOCSY (green) and NOESY (red) spectra, showing thymine Me-H6 assignments.
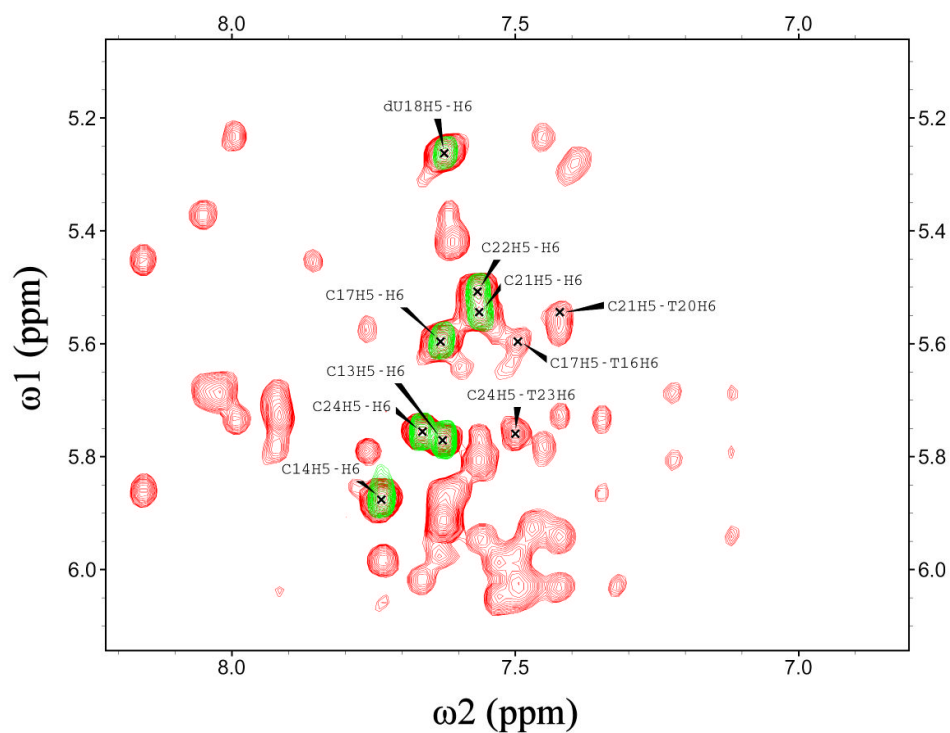
Figure 2.9. Expansions of the TOCSY (green) and NOESY (red) spectra, showing cytidine/deoxyuridine H5-H6 assignments.

H1′ assignments could then be made for each of the pyrimidine residues from correlations between H1′-H6 protons. These occur in the same region of the spectrum as the H5-H6 correlations, but can be easily distinguished due to the fact that H1′ resonances exhibit TOCSY connections to other sugar protons, and H1′ protons generally resonate at a slightly higher chemical shift than H5 protons (Figure 2.10).
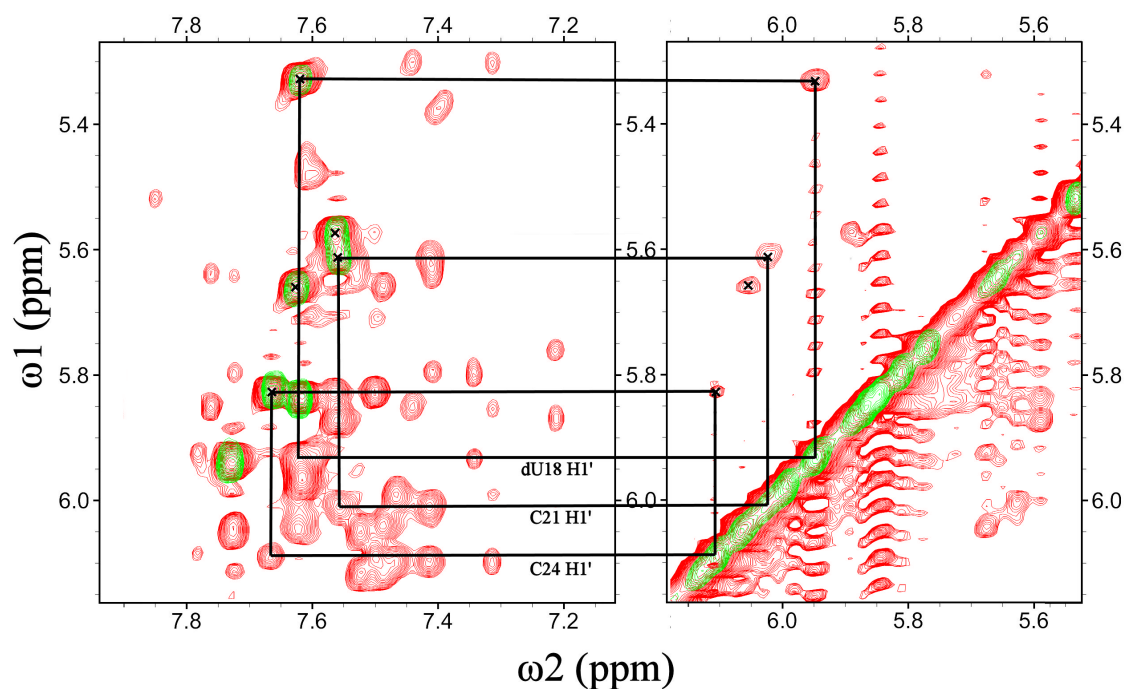
Figure 2.10. Expansions of the NOESY (red) and TOCSY (green) spectra, showing some H6-H5-H1′ connections.

The next stage in the assignment process was to identify the purine H1′-H8 cross peaks, which are present in the same region of the NOESY spectrum as the C/dU H5/H1′-H6 correlations. Expected H1′-H8 correlations are indicated in Figure 2.11. These couplings were more difficult to assign because in this hairpin structure, the guanosine (G) and adenosine (A) residues are not neighboured by any pyrimidine residues, therefore correlations to these cross peaks from any previously assigned

residues were not observed. It was however, possible to identify a full network of correlations corresponding to the twelve purine residues, with the aid of inter-residue H1′-H8 and H8-H8 nOes (Figures 2.12 and 2.13).
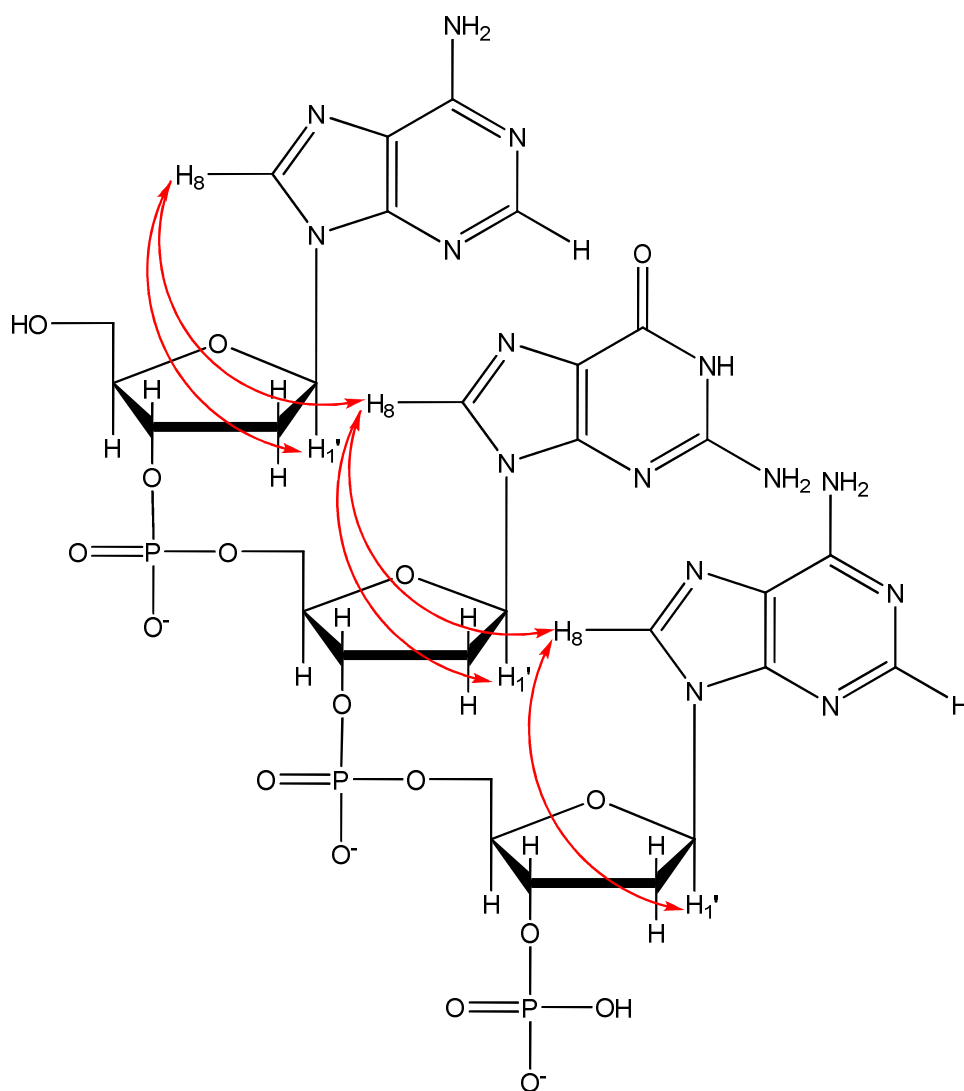
Figure 2.11. Expected NOESY connections between the H1′ and H8 protons of guanosine and adenosine nucleotides in a B-type DNA duplex.
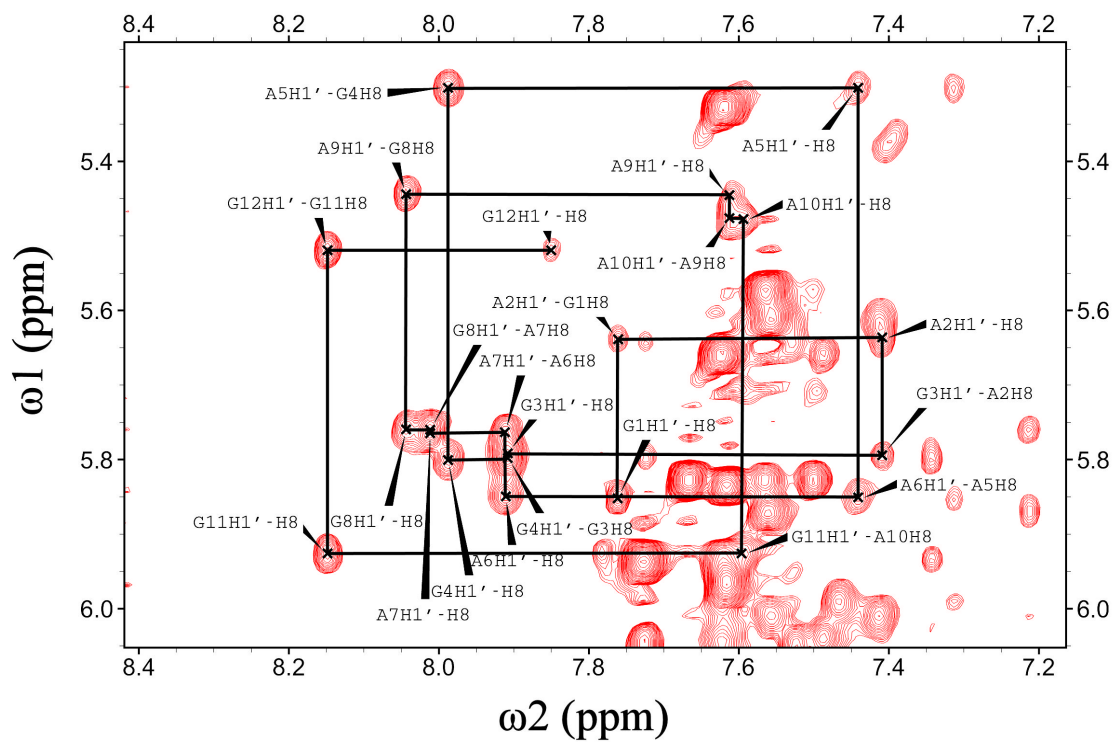
Figure 2.12. Expansion of the NOESY spectrum, showing H1′-H8 assignments. The nOe correlation pathway is highlighted.
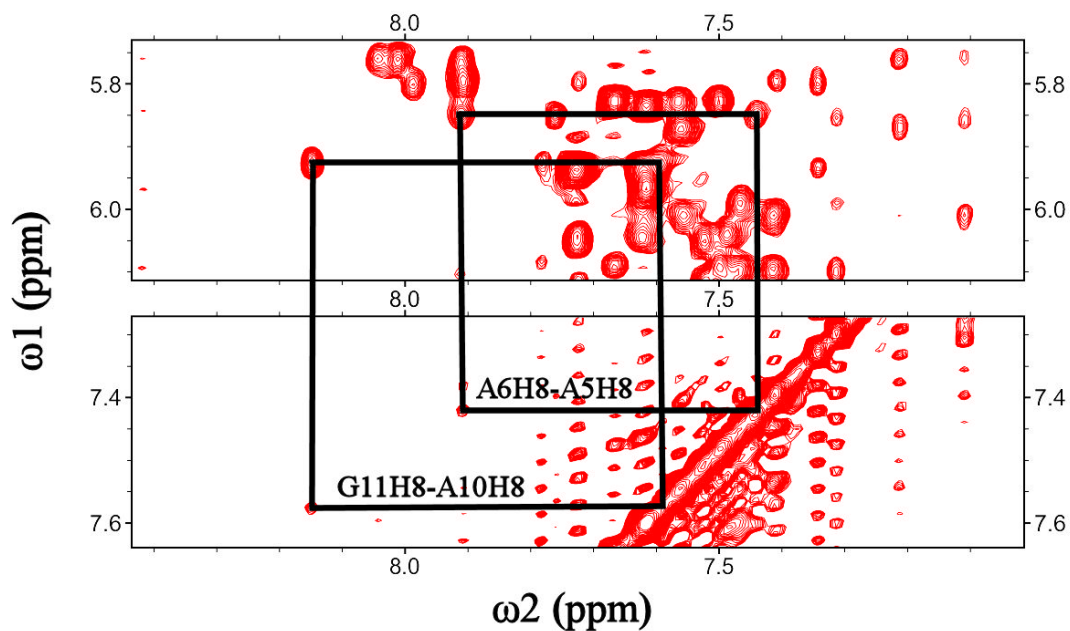
Figure 2.13. Expansion of the NOESY spectrum, showing some H8-H8 connections.

Once the base and H1′ protons had been assigned for every residue in the sequence, correlations to the remaining sugar protons completed the assignment process for the hairpin structure. These are highlighted in Figures 2.14 and 2.15. Several H2′/H2″-H1 inter-nucleoside cross peaks were observed (Figure 2.14) which provided further validation for the assignments made.
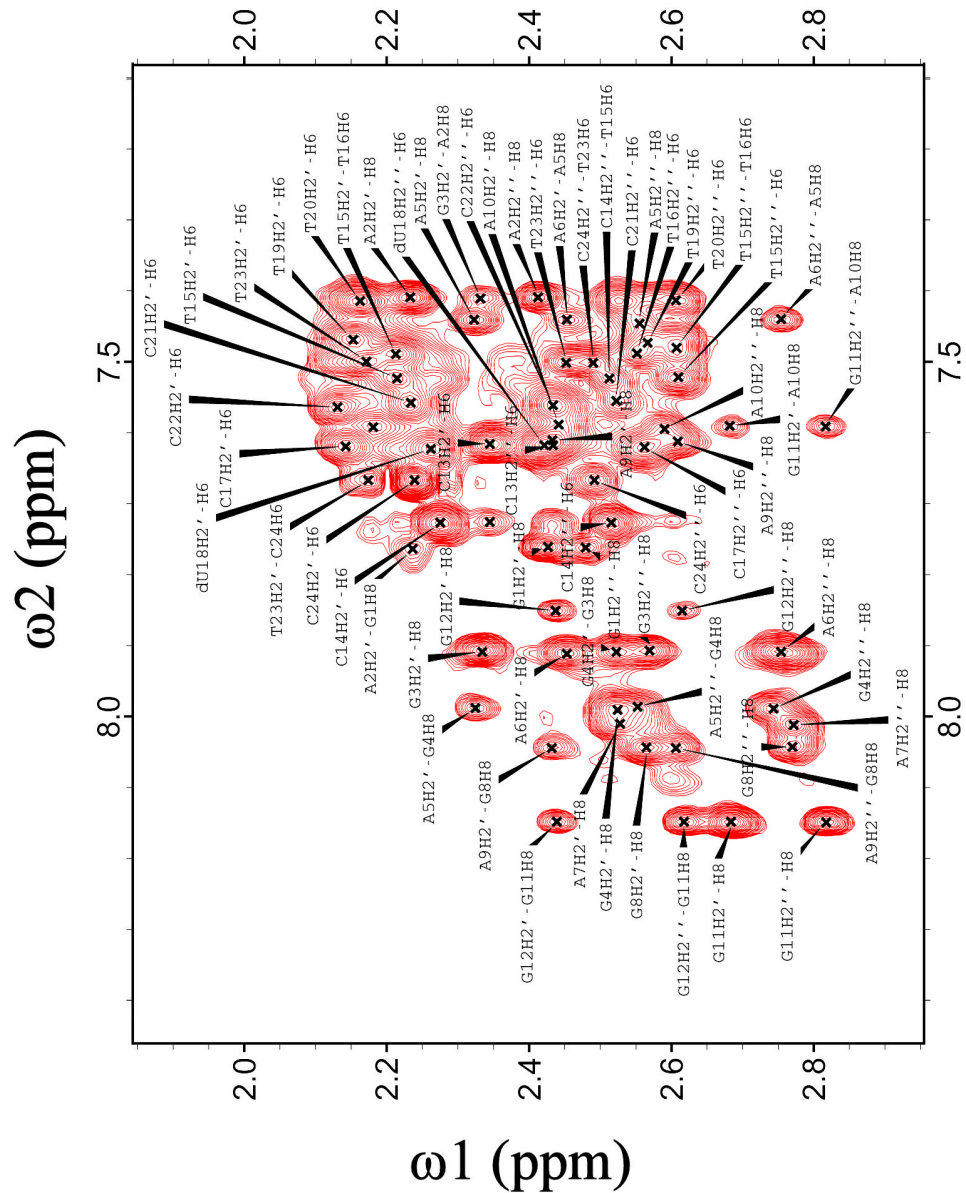
Figure 2.14. Expansion of the NOESY spectrum showing H2′/H2″-H6 correlations.
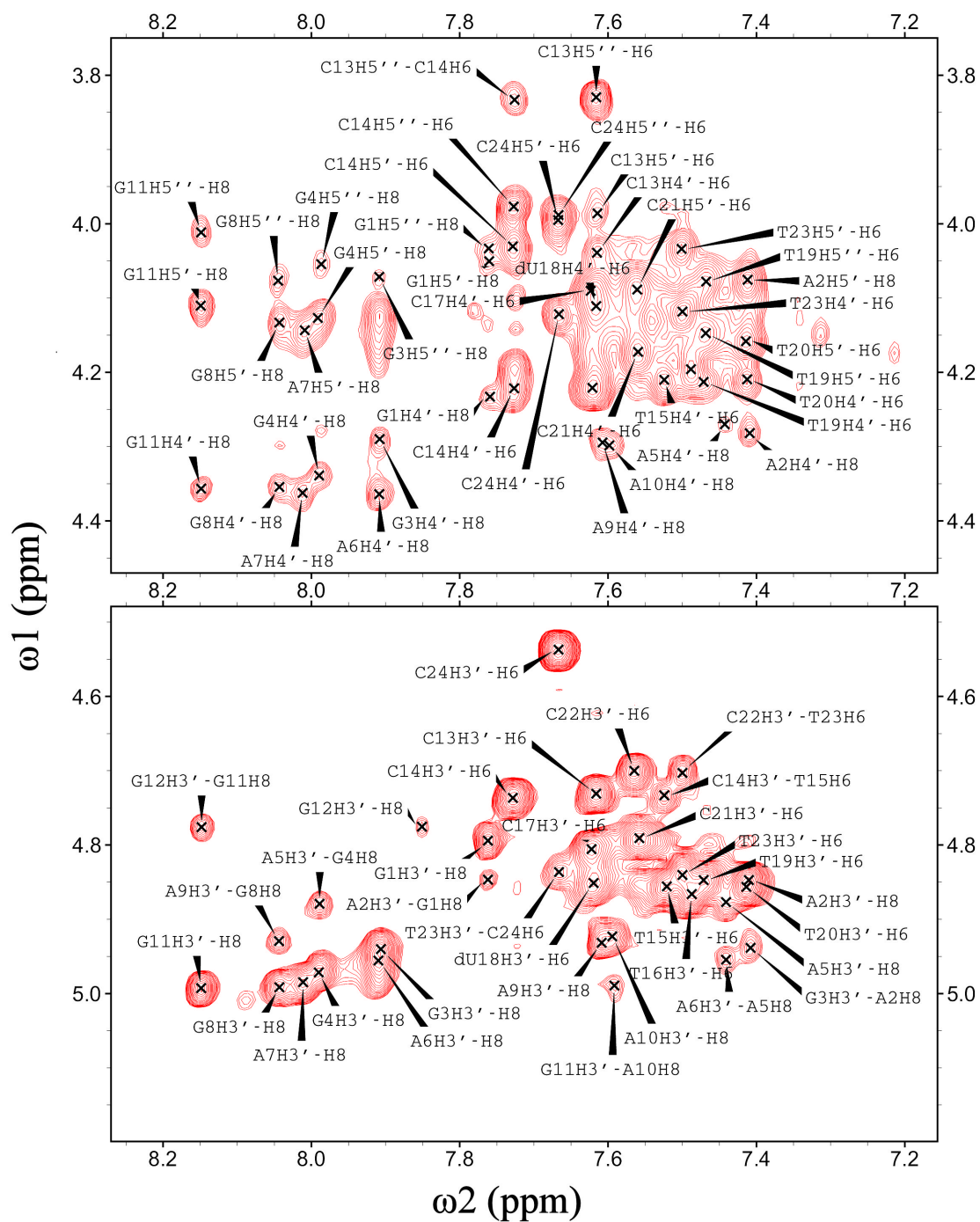
Figure 2.15. Expansions of two regions of the NOESY spectrum showing H4′/H5′/H5″-H6 and H3′-H6 correlations.

All non-exchangeable protons were assigned for the hairpin structure, with the exception of a few H5′/H5″ protons due to extensive overlap of these resonances. The chemical shifts are presented in Table 2.1.

Table 2.1. Chemical shifts (in ppm) of all non-exchangeable protons for the hairpin at 20 °C, referenced to H₂O at 4.8 ppm.

Residue	H8/H6	Me/H5	H1′	H2′	H2″	H3′	H4′	H5′/H5″
G1	7.76	-	5.85	2.43	2.48	4.80	4.23	4.05, 4.03
A2	7.41	-	5.64	2.23	2.41	4.85	4.28	4.08
G3	7.91	-	5.79	2.33	2.57	4.94	4.29	4.13, 4.07
G4	7.99	-	5.80	2.53	2.74	4.97	4.34	4.13, 4.06
A5	7.44	-	5.30	2.32	2.56	4.88	4.27	4.13, 4.06
A6	7.91	-	5.85	2.46	2.76	4.95	4.37	4.28, 4.24
A7	8.01	-	5.76	2.52	2.77	4.99	4.35	4.14, 4.07
G8	8.04	-	5.76	2.57	2.77	4.99	4.36	4.13, 4.08
A9	7.61	-	5.45	2.43	2.61	4.93	4.29	4.12, 4.08
A10	7.60	-	5.48	2.45	2.59	4.92	4.30	-
G11	8.15	-	5.93	2.68	2.82	4.99	4.36	4.11, 4.01
G12	7.85	-	5.56	2.44	2.62	4.78	4.36	4.11, 4.01
C13	7.62	5.84	5.97	2.34	2.43	4.73	4.04	3.99, 3.83
C14	7.73	5.94	6.05	2.51	2.27	4.74	4.22	4.04, 3.98
T15	7.52	1.66	6.12	2.21	2.51	4.86	4.21	4.17, 4.10
T16	7.49	1.66	6.04	2.22	2.55	4.86	4.20	-
C17	7.62	5.66	6.05	2.14	2.56	4.81	4.09	-
dU18	7.62	5.33	5.94	2.26	2.42	4.85	4.11	-
T19	7.47	1.61	5.99	2.14	2.57	4.85	4.21	4.15, 4.08
T20	7.41	1.61	6.10	2.16	2.61	4.86	4.21	4.16
C21	7.56	5.61	6.01	2.23	2.53	4.79	4.18	4.09
C22	7.57	5.57	5.83	2.13	2.44	4.70	4.13	4.03, 3.99
T23	7.50	1.68	6.09	2.18	2.45	4.84	4.12	4.03, 3.99
C24	7.67	5.83	6.25	2.24	2.50	4.54	4.12	3.99

2.3 Assignment of exchangeable proton resonances

Exchangeable or labile protons consist of the amino and imino base protons, which are shown in Figure 2.16. These protons cannot be detected in D_2O spectra as they exchange with deuterons in the solvent. Assignment of these protons therefore required a NOESY spectrum to be recorded for the hairpin in a 9:1 H_2O:D_2O solvent mix. This spectrum was recorded at 283 K in order to slow down the rate of exchange (Figure 2.17).
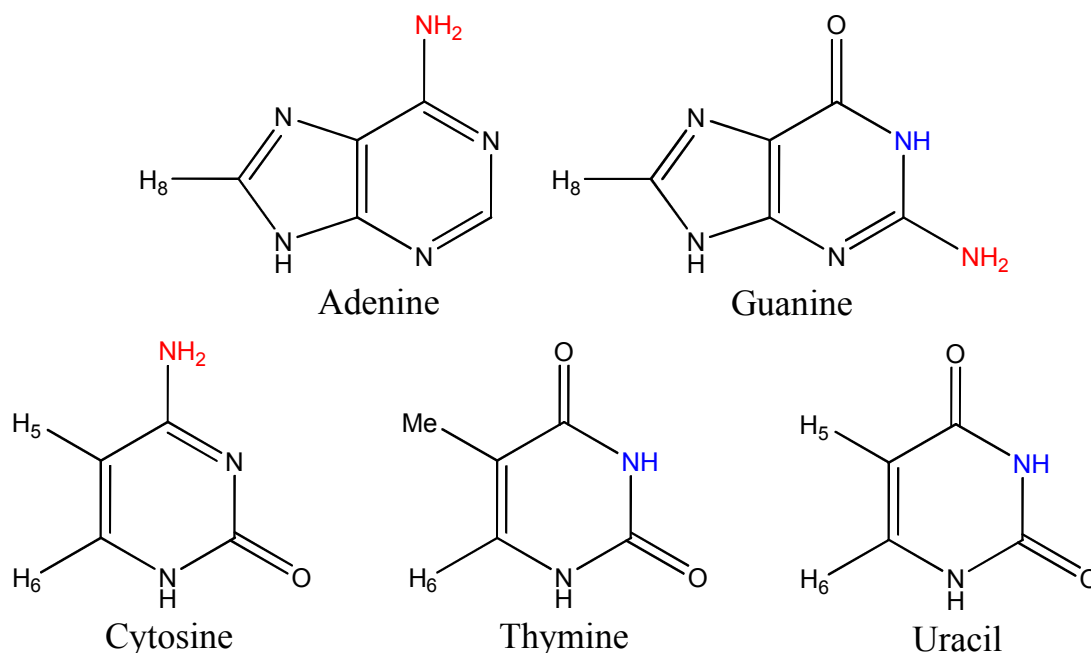
Figure 2.16. The structure of the heterocyclic base units, with the exchangeable protons highlighted. Imino protons are shown in blue, amino protons are labelled in red.
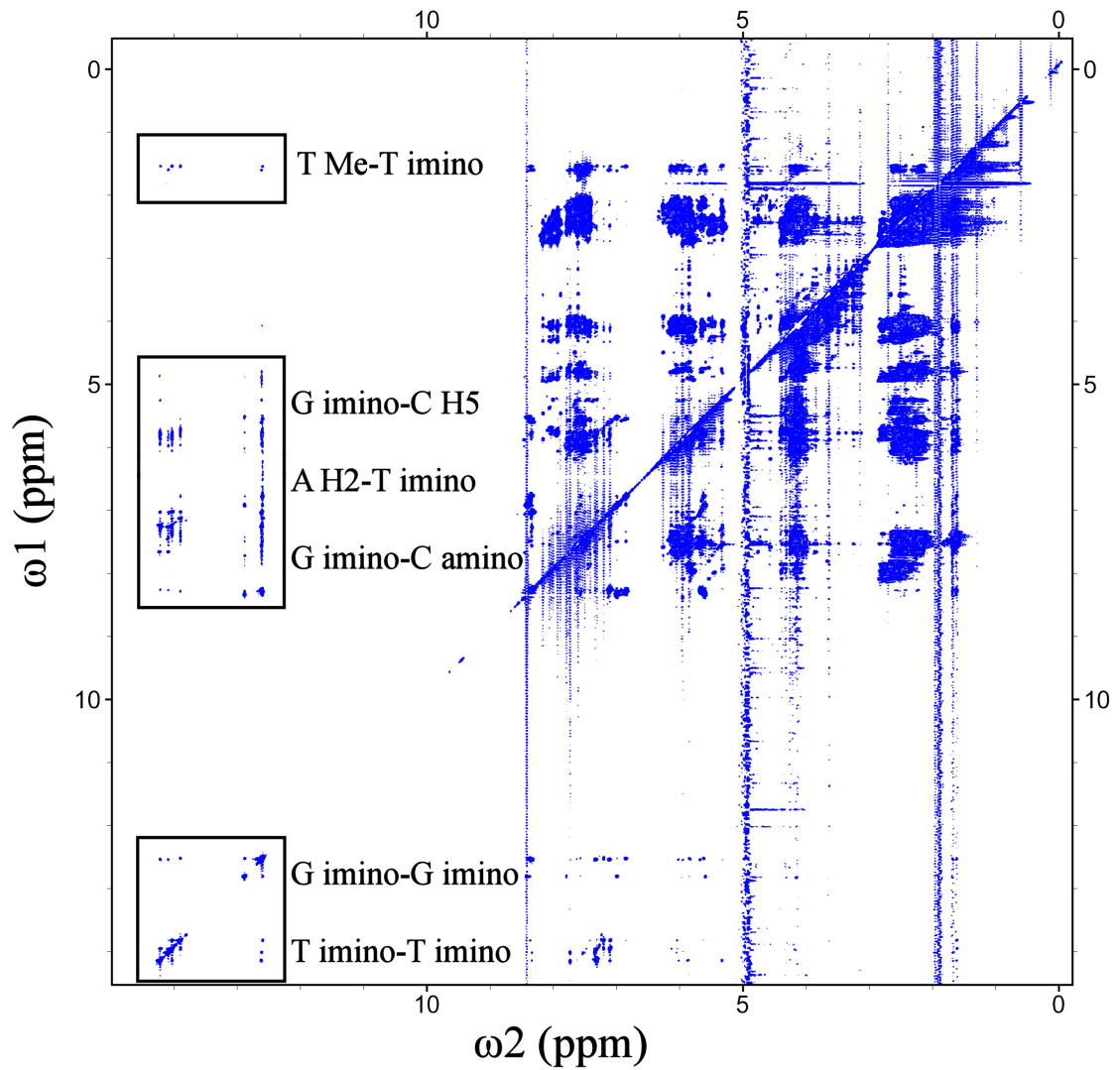
Figure 2.17. Full 750 MHz NOESY spectrum for the hairpin in a (9:1) H$_2$O:D$_2$O solution.

The imino and amino protons can yield information regarding the hydrogen bonding features and base pairing of an oligonucleotide[111] (Figure 2.18). Identification of the labile protons can also enable the adenosine H2 protons to be assigned, which do not exhibit many correlations to other non-exchangeables but may display nOe correlations to thymidine/deoxyuridine imino protons[112].
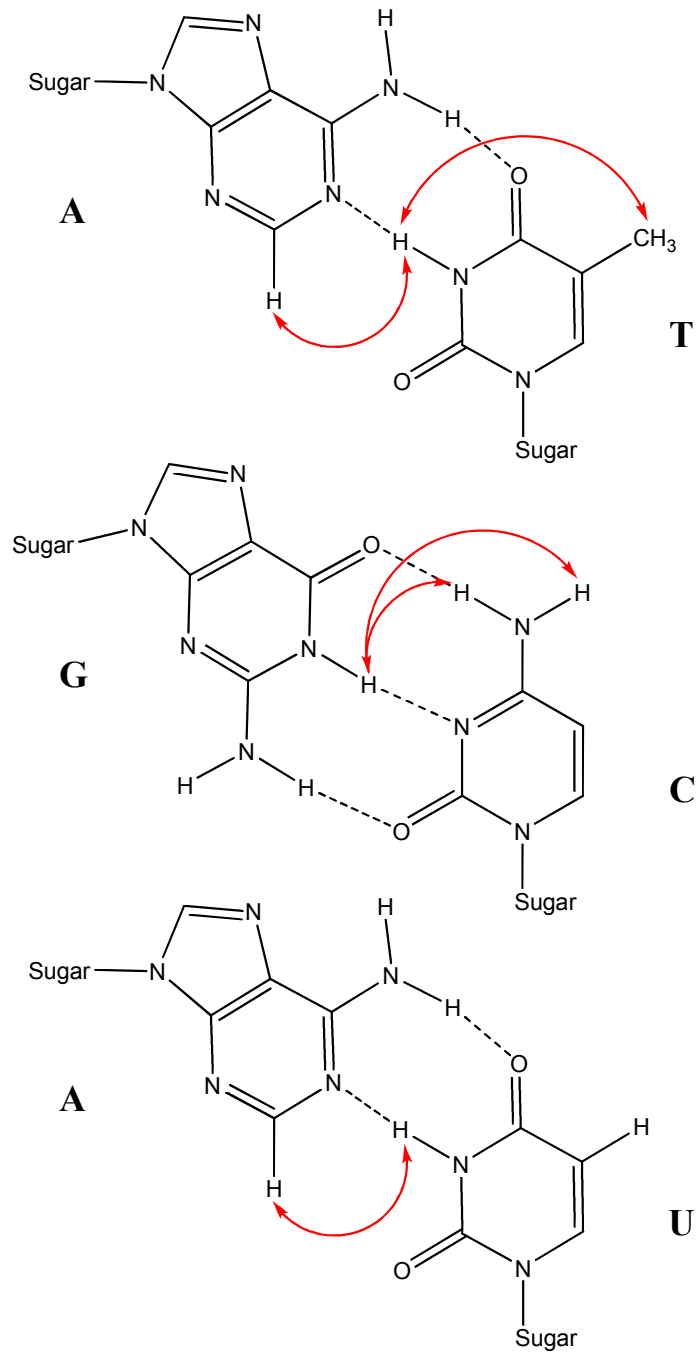
Figure 2.18. Expected nOe correlations for the labile protons in a Watson-Crick hydrogen bonded oligonucleotide.

Imino proton resonances can exhibit a correlation to imino protons on both adjacent residues and nearby cross-strand residues[111]. The expected imino couplings for the hairpin are highlighted in Figure 2.19.

$$G_1 \quad A_2 \quad G_3 \leftrightarrow G_4 \quad A_5 \quad A_6 \quad A_7 \quad G_8 \quad A_9 \quad A_{10} \quad G_{11} \leftrightarrow G_{12}$$
$$C_{24} \quad T_{23} \quad C_{22} \quad C_{21} \quad T_{20} \leftrightarrow T_{19} \leftrightarrow dU_{18} \quad C_{17} \quad T_{16} \leftrightarrow T_{15} \quad C_{14} \quad C_{13}$$
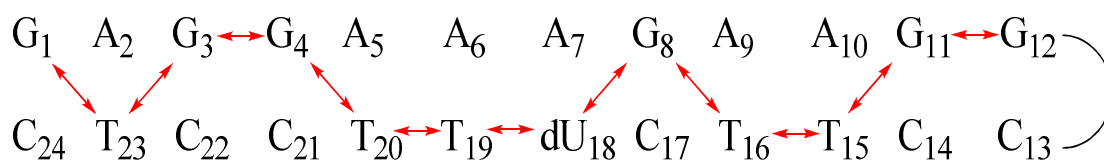
Figure 2.19. Expected nOe connections between G/T imino protons for the hairpin structure.

Imino protons were assigned via correlations to the methyl group of neighbouring T residues, and from intra- and inter-residue imino-imino couplings. A network of correlations between the methyl and imino protons can be mapped out and is shown in Figure 2.20.
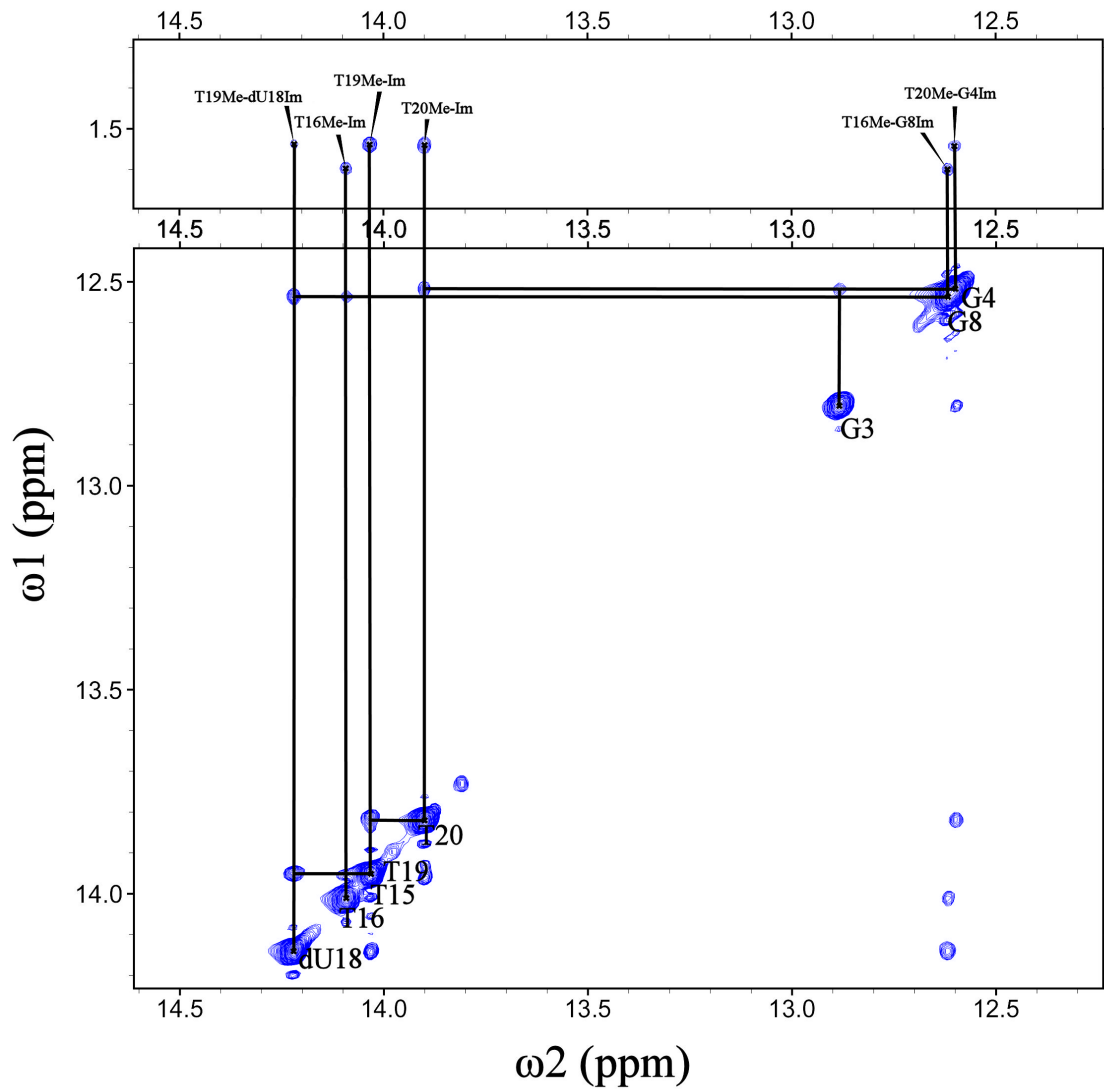
Figure 2.20. Expansions of the NOESY spectrum for the hairpin in a (9:1) H₂O:D₂O solution, showing Me-imino and imino-imino correlations.

Cytidine amino protons, both hydrogen bonded and non-hydrogen bonded could then be assigned from correlations to the G imino protons (Figure 2.21) and A H2 protons were identified from cross-strand couplings to T imino protons (Figure 2.22).
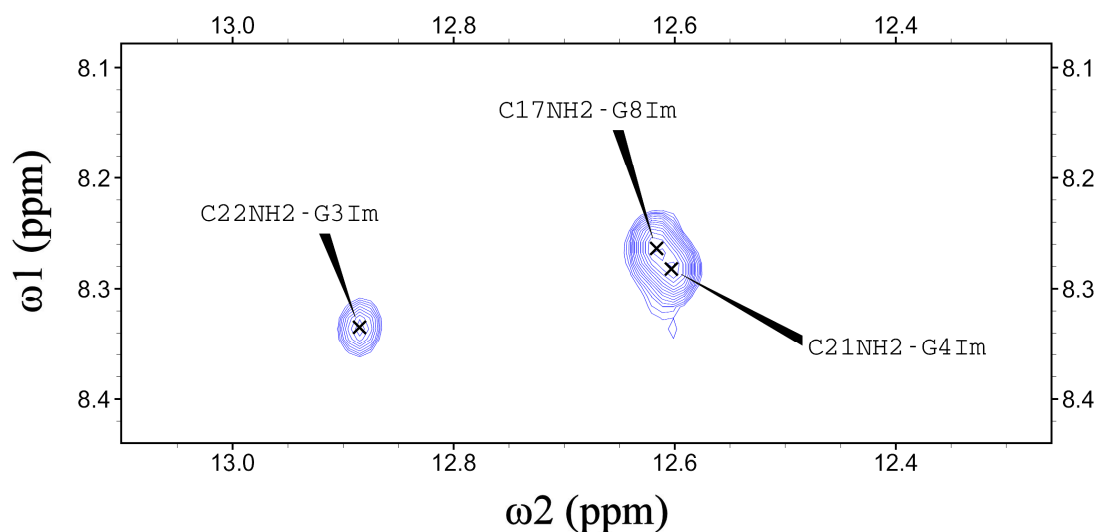
Figure 2.21. Expansion of the NOESY spectrum of the hairpin in a (9:1) H₂O:D₂O solution, showing the region in which the hydrogen bonded C amino-G imino couplings are found. Non-hydrogen bonded amino protons are shifted approx. 1-1.5 ppm upfield.
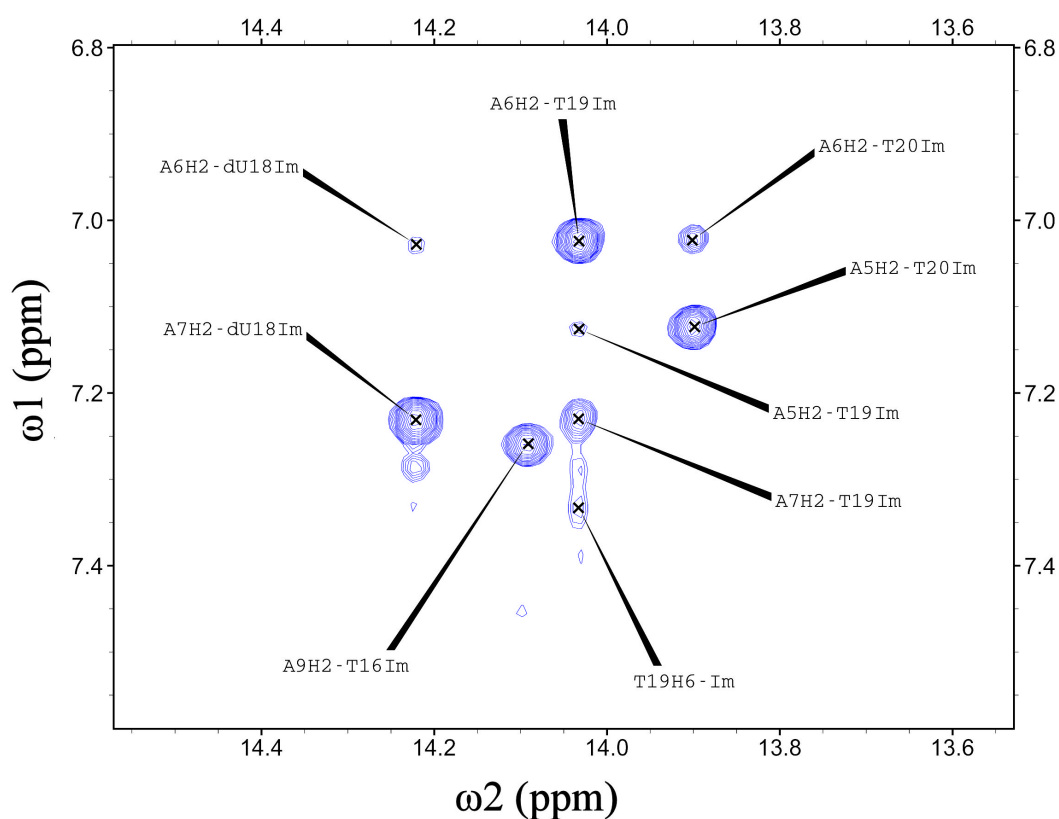
Figure 2.22. Expansion of the NOESY spectrum of the hairpin in a (9:1) H₂O:D₂O solution showing A H2-T imino couplings.

Exchangeable protons were assigned for all residues except those at the ends of the oligonucleotide (G1, G11, G12 and T23), which is most likely due to fraying at either end of the hairpin, leading to an increased exchange rate with the solvent due to increased accessibility at the ends. The chemical shifts of the assigned labile and H2 protons are given in Table 2.2.

Table 2.2. Chemical shifts (in ppm) of all exchangeable and H2 protons for the hairpin at 2 °C, referenced to H_2O at 4.95 ppm.

Residue	Imino	Amino	H2
G1	-	-	-
A2	-	-	-
G3	12.88	-	-
G4	12.60	-	-
A5	-	-	7.13
A6	-	-	7.02
A7	-	-	7.24
G8	12.62	-	-
A9	-	-	7.26
A10	-	-	-
G11	-	-	-
G12	-	-	-
C13	-	-	-
C14	-	-	-
T15	13.95	-	-
T16	14.09	-	-
C17	-	8.27	-
dU18	14.22	-	-
T19	13.89	-	-
T20	14.03	-	-
C21	-	8.29	-
C22	-	8.34	-
T23	-	-	-
C24	-	-	-

2.4 Qualitative structural analysis

The sequential nOe connections that were observed for the hairpin indicate that the structure adopts the B-type helix that is expected for DNA structures. However, further qualitative analysis is required in order to determine local features of the structure and provide a more precise model. Cross peak height and volume measurements taken from NOESY spectra indicate how close in space two protons are, which can provide valuable structural information regarding the conformation of individual nucleotides such as the sugar pucker and glycosidic bond orientation.

2.4.1 Sugar Pucker Determination

The type of helix formed in a nucleic acid is defined by the conformation of the ribose or deoxyribose sugars. As mentioned in Chapter 1, in the A-type helix (typical of RNA) , sugars have a preference for the C3′-*endo*, or north pucker, whereas B-type helical structures (characteristic of DNA) contain sugars which tend to adopt the C2′-*endo*, or south pucker[1] (Figure 1.4).

The conformation of each individual sugar can be determined by H2′-H6/H8 and H1′-H2′/H2″ nOe measurements. The NOESY cross peak volumes for these protons are inversely proportional to the conformation dependent distances between the protons, whereby the greater the cross peak volume, the closer in space the two protons are[109].

Figure 2.23 indicates the relative H2′-H6/H8 distances for each of the sugar puckers. In the north conformation, the H2′ proton of one residue is closer to the H6 of the preceding 5′ base than of its own H6 proton, whereas the opposite is true for the

south pucker. Therefore, for B-type DNA structures with all south sugar puckers, the H2′(n)-H6/H8(n) cross peaks would have a larger volume than the inter-residue H2′(n)-H6/H8(n-1) cross peaks.

North puckers   South puckers
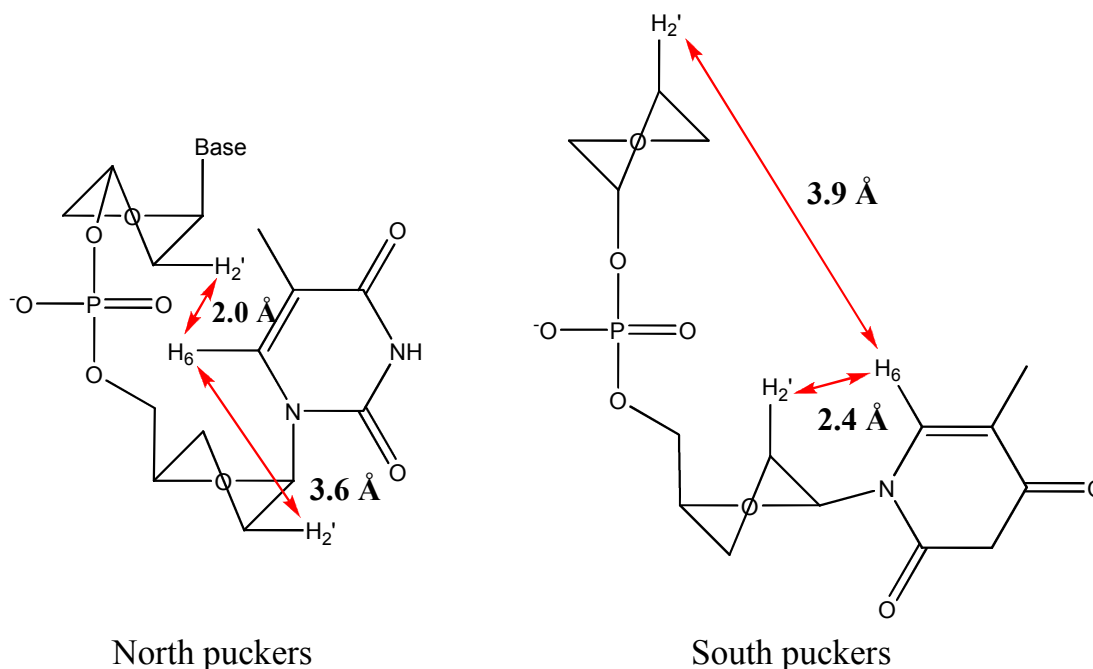
Figure 2.23. H2′-H6/H8 distances in north and south sugar puckers[113].

Due to extensive overlap in the H2′-H6/H8 region of the NOESY spectrum, it is often difficult to accurately measure peak volumes and a more appropriate approach is to measure peak heights, which can be obtained more easily and which have previously been used to good effect[114]. The height of the H2′-H6/H8 nOe peaks were therefore measured and compared to determine the sugar pucker adopted by each residue (peak heights for residue G11 are shown in Figure 2.24). Table 2.3 gives the sugar pucker for each nucleotide in the sequence for which peak heights could be obtained. Several residues could not be analysed due to an absence of peaks or extensive overlap. All of the residues that were analysed using the H2′-H6/H8 peak heights were found to adopt the south sugar pucker preferentially.
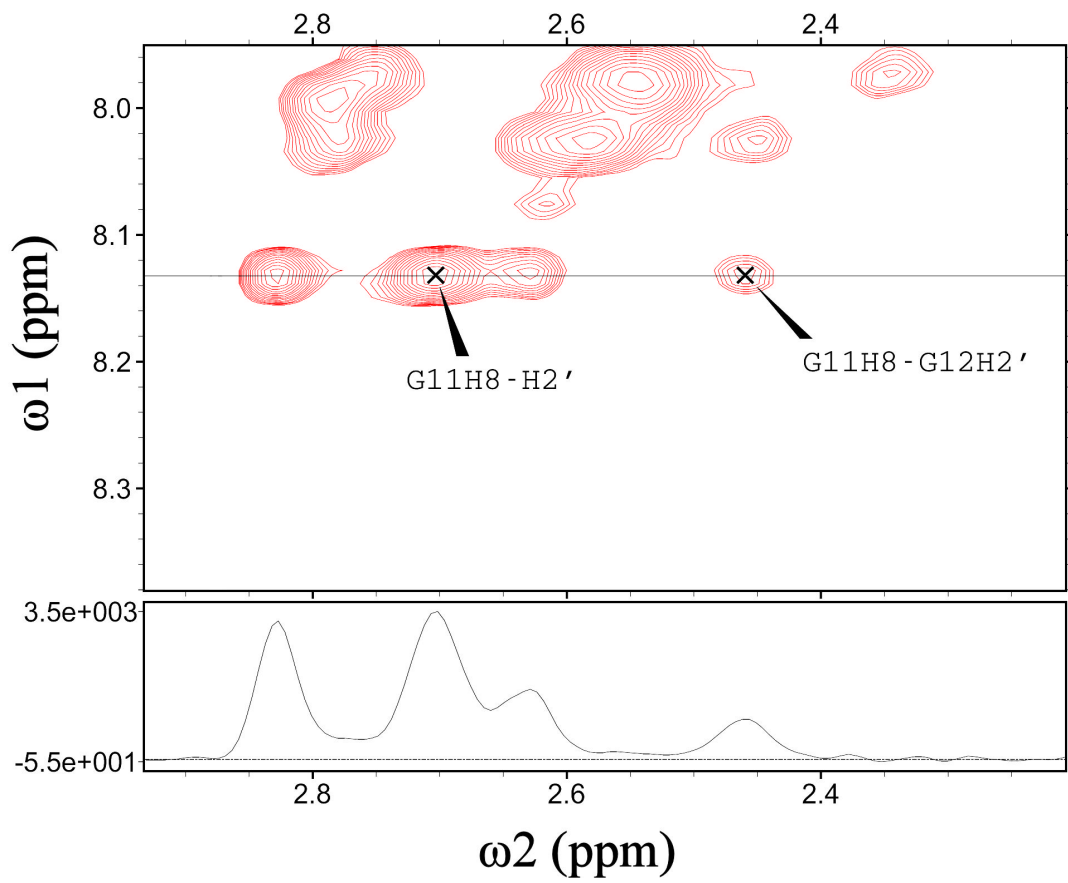
Figure 2.24. One dimensional slice through the hairpin NOESY spectrum showing relative heights of the G11 H2′-H8 and G12 H2'-G11 H8 cross peaks.

Table 2.3. Sugar puckers of the hairpin residues deduced from the relative peak heights of intra- and inter-residue H2′-H6/H8 NOESY cross peaks.

Residue	Sugar pucker	Residue	Sugar Pucker
G1	South	C13	-
A2	South	C14	-
G3	South	T15	-
G4	South	T16	-
A5	South	C17	South
A6	-	dU18	-
A7	-	T19	South
G8	South	T20	South
A9	South	C21	-
A10	South	C22	-
G11	South	T23	South
G12	-	C24	South

Note: South pucker when H2′(n)-H6(n) > H2′(n)-H6(n+1)

H2′(n)-H8(n) > H2′(n+1)-H8(n)

North pucker when H2′(n)-H6(n) < H2′(n)-H6(n+1)

H2′(n)-H8(n) < H2′(n+1)-H8(n)

- Peak height could not be obtained due to overlap

The distances between the H1′ and H2′/H2″ protons are also conformation dependent (Figure 2.25). In the south pucker, the H1′ of one residue is much closer in space to its H2″ than to its H2′, so the H1′-H2″ cross peak will have a larger volume/height than the H1′-H2′ peak. Again the opposite is true for the north sugar pucker.
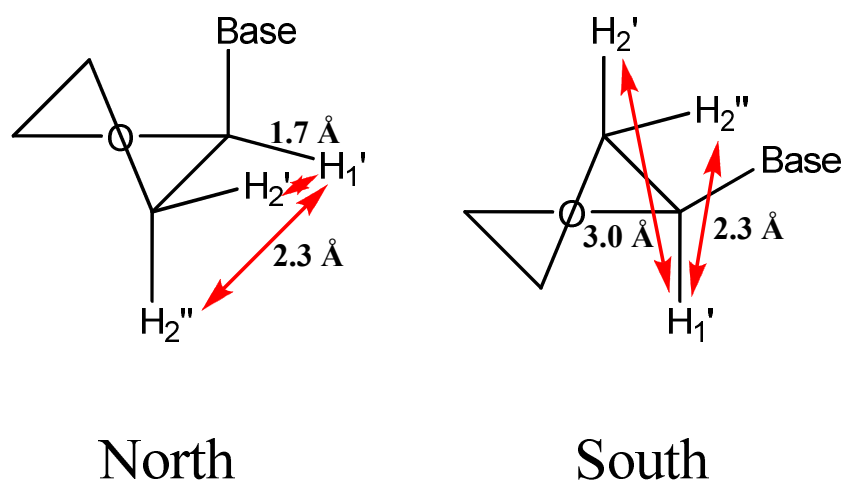
Figure 2.25. H1′-H2′/H2″ distances in north and south sugar puckers[113].

As the H1′-H2′/H2″ region of the NOESY spectrum is fairly crowded with several overlapping peaks, again the heights rather than the volumes of the peaks were examined. Figure 2.26 shows the relative H1′-H2′/H2″ peak heights for residue A5. The sugar puckers for each nucleotide in the sequence for which peak heights could be obtained are given in Table 2.4. As before, several residues could not be characterised due to an absence of peaks or extensive overlap. All of the residues that were analysed using this method were found to contain south sugar puckers.
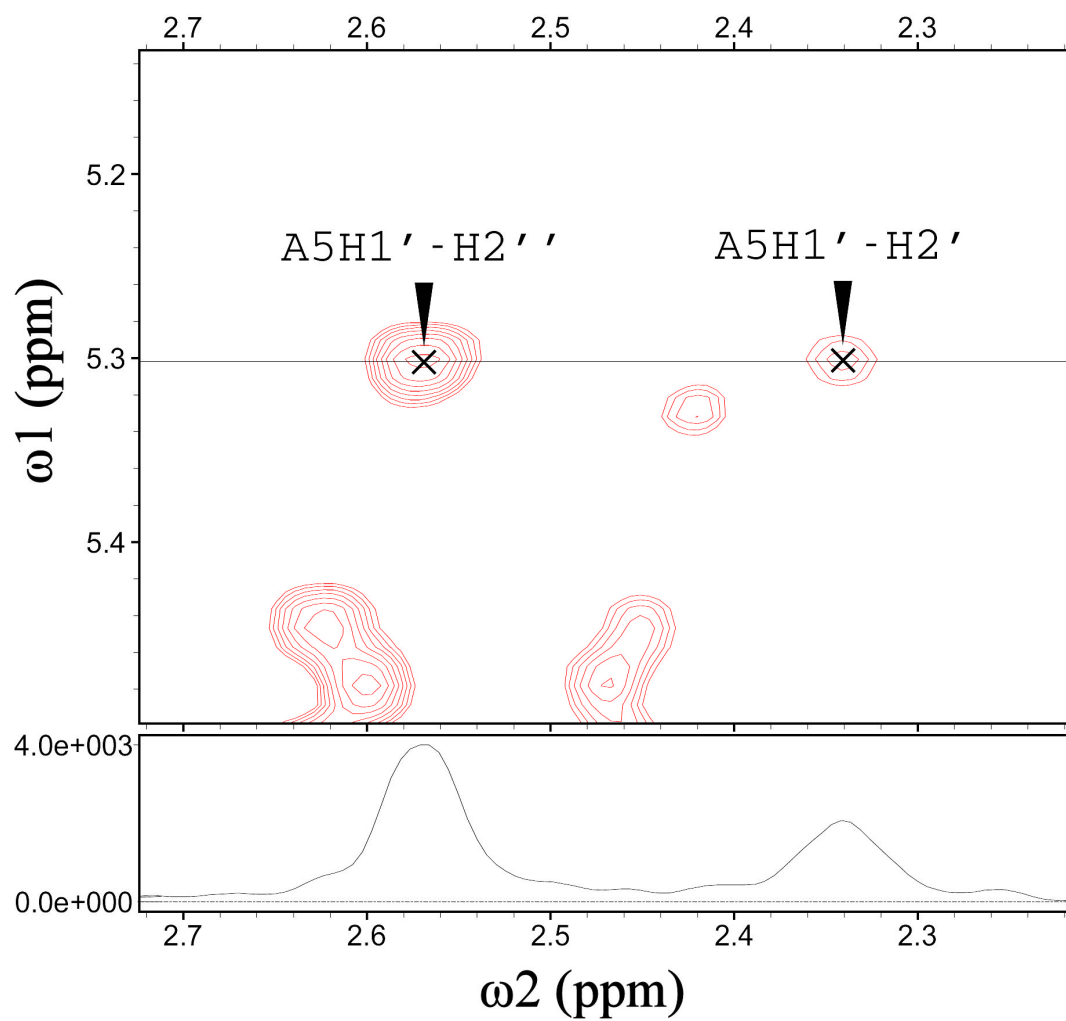
Figure 2.26. One dimensional slice through the hairpin NOESY spectrum showing relative heights of the A5 H1'-H2' and H1'-H2'' cross peaks.

Table 2.4. Sugar puckers of the hairpin residues deduced from the relative cross peak heights of H1′-H2′/H2″ NOESY cross peaks.

Residue	Sugar pucker	Residue	Sugar Pucker
G1	South	C13	South
A2	South	C14	South
G3	South	T15	South
G4	South	T16	South
A5	South	C17	-
A6	-	dU18	South
A7	South	T19	South
G8	South	T20	-
A9	South	C21	South
A10	South	C22	South
G11	South	T23	-
G12	South	C24	-

Note: South pucker when H1′-H2′ > H1′-H2″

North pucker when H1′-H2′ > H1′-H2″

-      Peak height could not be obtained due to overlap

All of the sugar puckers within the hairpin could be determined by measurement of the H2′-H6/H8 and/or H1′-H2′/H2″ cross peaks with the exception of A6. All other residues were found to adopt the south sugar pucker, which is typical of a standard B-type DNA duplex, suggesting that the general structure of the hairpin will be that of a B-type helix.

2.4.2 Glycosidic Bond Orientation

The base can adopt two main orientations relative to the sugar ring, *syn* and *anti*, which are dictated by the glycosidic torsion angle (O4′-C1′-N9-C4 in purines, O4′-

C1′-N1-C2 in pyrimidines[1]). The *anti* conformation is present for both A- and B-type helices, with the sugar in the C3′-*endo* (north) pucker for RNA, and C2′-*endo* (south) for DNA[8].

The orientation can be determined by nOe measurements between the H6/H8 and H1′/H2′/H3′ protons. The *anti* conformation is preferred when there is a strong nOe correlation (short proton-proton distance) between the H6/H8 of the base and H2′ or H3′ of the sugar, while a strong nOe correlation between H6/H8 and H1′ indicates the *syn* conformation is adopted preferentially[115] (Figure 2.27).

Syn

Anti (south)

Anti (north)
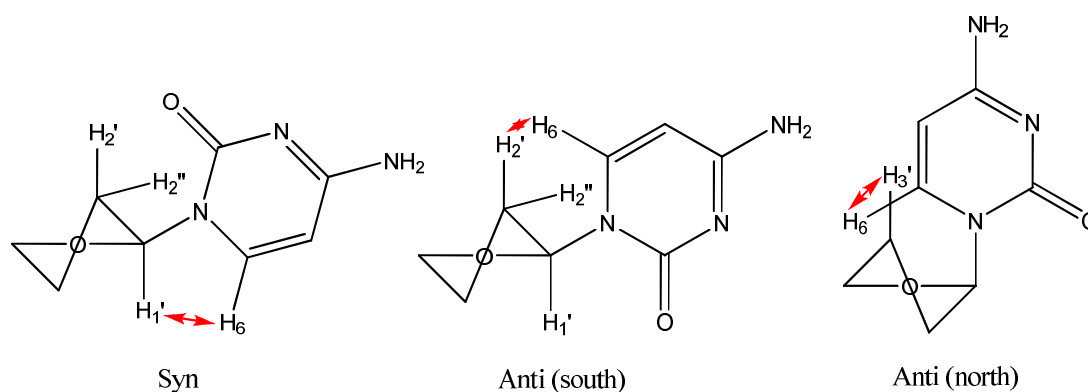
Figure 2.27. Glycosydic bond orientations relative to the sugar puckering, showing the strongest nOe correlations expected for each orientation.

As was true for the assessment of sugar puckers via nOe measurements, the H1′/H2′/H3′-H6/H8 cross peaks resonate in a fairly crowded region of the spectrum, so it was more appropriate to measure peak heights as opposed to peak volumes. Table 2.5 shows the glycosidic bond orientation for each residue of the hairpin.

Table 2.5. Glycosidic bond orientation of the hairpin residues deduced from the relative heights of H1′/H2′/H3′-H6/H8 NOESY cross peaks.

Residue	Orientation	Residue	Orientation
G1	Anti-S	C13	Syn
A2	Anti-S	C14	Anti-S
G3	Anti-S	T15	Anti-S
G4	Anti-S	T16	-
A5	Anti-S	C17	Anti-N
A6	Anti-S	dU18	Anti-S
A7	Anti-S	T19	Anti-S
G8	Anti-S	T20	Anti-S
A9	Anti-S	C21	Anti-S
A10	Anti-S	C22	Anti-S
G11	Anti-S	T23	Anti-S
G12	Anti-S	C24	Anti-S

Note: Syn orientation when H1′-H6/H8 > H2′/H3′-H6/H8

Anti-S orientation when H2′-H6/H8 > H1′/H3′-H6/H8

Anti-N orientation when H3′-H6/H8 > H2′/H1′-H6/H8

-        Peak height could not be obtained due to overlap

The majority of nucleotides were found to adopt an *anti* orientation with a south sugar pucker, as expected for a standard B-type helix.

2.5 Discussion

The target hairpin was successfully assigned using the procedure for a typical B-type DNA structure. The majority of the non-exchangeable protons were assigned with the exception of some H5′/H5″ protons, which is usual for an oligonucleotide of this

size[116, 117]. The NOESY correlations that were observed are typical of those seen for a B-type DNA double helix, indicating that the assumed B-type model was correct. Assignment of the labile protons was achieved for all residues except those at the two ends of the hairpin, due to end-fraying and solvent exchange. Observation of the hydrogen bonded imino and amino protons provided further evidence of a double helical structure with Watson-Crick base pairing[110].

Qualitative analysis was performed for most of the residues in the hairpin, which allowed localised structural features to be determined. The majority of nucleotides analysed were found to preferentially adopt the south sugar pucker and *anti* glycosidic bond orientation which is typical for residues in a B-type DNA double helix. This is in agreement with the hairpin NMR assignments and suggests that the hairpin retains a B-type structure throughout the length of the structure, with no unusual local features. However, for a more in-depth analysis of the structure of the target hairpin, quantitative analysis is required and is detailed in the following chapter.

2.6 Conclusions

The target hairpin was successfully assigned by NMR using a B-type model. The oligonucleotide was shown to adopt a B-type double helix which participates in Watson-Crick hydrogen bonding. Qualitative analysis indicated that the hairpin contains no unusual localised structural features and that the majority of residues adopt a south sugar pucker and *anti* glycosidic bond orientation.

Chapter 3

NMR Structure Generation for the Hairpin

Chapter 3.   Structure Generation for the Hairpin

3.1 Introduction

A qualitative structural analysis of the target hairpin duplex was performed and described in Chapter 2. However, a full, and quantitative structure determination is required for a more in depth analysis of the local and global structural features of the oligonucleotide. This will also be of great interest for comparison to the structures of both the native and modified triplexes, (detailed in Chapter 5). The molecular modelling program, Crystallography and NMR Systems (CNS) Solve[118], will be used to generate low energy structures for the hairpin duplex, using nOe-derived proton-proton distances as structural constraints.

3.2 Structure determination by NMR

Structure generation using experimentally determined NMR constraints requires, in the first instance, the assignment of resonances in NOESY spectra from which inter-proton distances may be calculated[119]. Inter-proton distances are inversely proportional to the volumes of the nOe cross peaks, and so peak volumes will increase with decreasing separation of two protons. These distances are used as experimental constraints, along with other data such as hydrogen bonding, dihedral angle and planarity constraints, to generate energy minimised, 3D structures. The general scheme for structure generation is illustrated in Figure 3.1.

Assignment of cross peaks in NOESY spectra
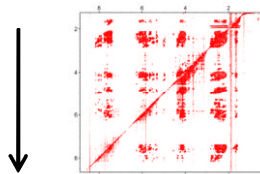
Calculation of NMR derived distance restraints

$$r_{ab} = r_{ref} \left(\frac{\eta_{ref}}{\eta_{ab}} \right)^{1/6}$$

Molecular Modelling
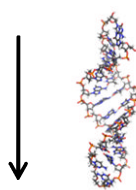
Energy Minimised 3D structure

Figure 3.1. Flow chart illustrating the general scheme for structure generation.

3.2.1 Calculation of Distance Restraints

There are several methods which can be used to extract inter-proton distances from 2D NOESY spectra, the most commonly employed being the isolated spin-pair approximation (ISPA)[120].

As mentioned, the intensities of the cross peaks in a 2D NOESY spectrum are related to the distances between protons which are in close spatial proximity (less than ~ 6 Å apart[109]). The volume of a cross peak is proportional to both the rate of magnetisation transfer between the two spins and the mixing time - the length of time in which the magnetisation is allowed to evolve[111]. At short mixing times (up to ~ 250 ms), the ISPA model assumes that the nuclei act as an isolated pair and the rate of

magnetisation transfer is dependent on the internuclear separation between the spins alone, with a distance dependency of r^{-6}. However, at longer mixing times, the distance dependence deviates from linearity due to multispin effects referred to as 'spin diffusion'[119]. Spin diffusion is the indirect magnetisation transfer *via* intervening spins, which as it becomes more prevalent, will eventually cause the NOESY cross peak intensities to become independent of distance[119].

The impact of spin diffusion can be assessed by constructing nOe build-up curves, i.e. cross peak volumes as a function of mixing time (Figure 3.2).
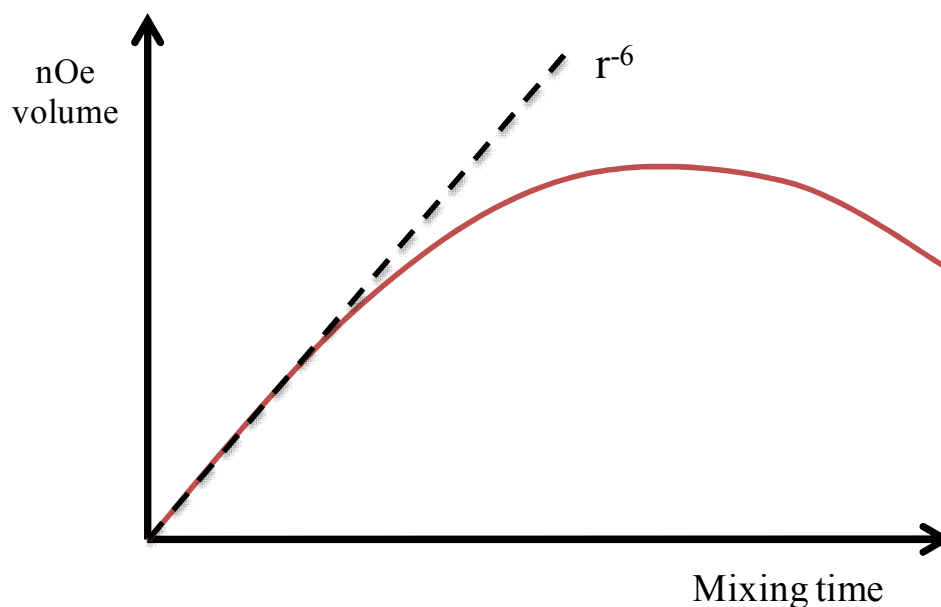
Figure 3.2. A build-up curve, indicating the linear region (dashed line) in which spin diffusion effects are minimal.

Spin diffusion can be considered negligible during the early, linear region of the plot; at these mixing times the ISPA can be used to calculate unknown distances from their relaxation rates[119] (Equation 3.1).

$$r_{ab} = r_{ref} \left(\frac{\eta_{ref}}{\eta_{ab}}\right)^{1/6} \qquad \textbf{Equation 3.1}$$

The ISPA uses a known, covalently fixed distance (r_{ref}) and volume (η_{ref}) as a reference, to calculate an unknown distance (r_{ab}) from its own nOe cross peak volume (η_{ab})[119]. This model, although the most convenient and commonly used, can result in large systematic errors due to not accounting for spin diffusion[121]. Analysis may also be limited, particularly at short mixing times in which the cross peaks have a lower signal-to-noise ratio, rendering some nOes undetectable[119].

Other, more complicated methods may provide an improvement on the errors associated with the ISPA model. For example, the complete relaxation matrix analysis (CORMA) takes all dipole-dipole interactions into account, enabling more accurate distances to be calculated[122]. A matrix of all magnetisation transfer pathways within a molecule is produced, and this is used to calculate proton-proton distances[123]. In this way, the whole proton network is taken into consideration and spin diffusion is accounted for[109]. An intensity matrix (I) is created from NOESY cross peak volumes, which is related to the complete dipole-dipole relaxation rate matrix (R) at a given mixing time τ_m, by Equation 3.2[123],

$$I(\tau_m) = e^{-R\tau_m} \qquad \textbf{Equation 3.2}$$

Off diagonal elements of R represent the cross-relaxation rates between two protons, and is dependent on molecular motions in addition to the distance between them[123].

However, a complete relaxation matrix cannot be constructed if assignments are incomplete, which is often the case due to peak overlap and unobservable nOes (for protons separated by > 6 Å). An incomplete matrix can lead to significant errors in distances, requiring the need for a 'hybrid intensity matrix', which uses an idealised data set for A- or B-type DNA[111]. Experimental distances calculated using the ISPA model are substituted into the theoretical spectrum and a complete matrix is constructed by back-calculating distances until there is a good agreement between the theoretical and experimental spectrum. The accuracy of this method can be determined by recording NOESY spectra with different mixing times, and producing a variety of intensity matrices. If the model has correctly accounted for spin diffusion effects, all distances calculated should be equal, regardless of the mixing time[123].

The programme MARDIGRAS (Matrix Analysis of Relaxation for Discerning the Geometry of an Aqueous Structure) adopts a similar approach to CORMA, as it also utilises an idealised intensity matrix combined with experimental cross peak volumes. The intensity matrix is converted to a relaxation rate matrix, which is then improved by an iterative procedure and distances are calculated from the final cross-relaxation rates[124].

Relaxation matrix approaches do tend to provide more accurate values than ISPA for shorter distances, however these methods are still subject to the same errors caused by poor signal:noise, errors in peak integration, short relaxation delays and internal motions[109]. Incorrect data is difficult to identify when using these complicated relaxation matrix methods, therefore a semi-quantitative approach to the more simplistic ISPA model is most widely used in calculating nOe distance restraints.

The semi-quantitative ISPA model relies on a large number of loose nOe constraints to restrain the conformation of a structure, ideally between 15-20 per residue[125]. Inter-proton distances can be calculated by integrating nOe cross peak volumes measured at a single, short mixing time which is known to be in the linear growth region and then inputting the values to Equation 3.1. Alternatively, a series of spectra can be recorded over a range of mixing times and relaxation rates extracted from an nOe build-up plot[111].

3.3 Molecular modelling

The ultimate aim of all molecular dynamics (MD) calculations is to optimise the agreement of an atomic model with experimental data and known chemical information. Molecular dynamics relies on a simplified system in which the atomic composition of a molecule is considered to be a collection of masses interacting with each other via harmonic forces[126]. These forces can be described by potential energy functions in terms of various structural features such as bond lengths, torsion angles and non-bonded interactions. A combination of these potential energy functions is known as a *force field*, the design and implementation of which is highly important in producing accurate structures[126,127].

MD simulations work on the basis that they explore the conformational space of the molecule to define all structures that are consistent with the experimental restraints. If the resulting structures are closely related and the molecule has been well defined by experimental data, it can be concluded that the structure has been successfully determined[119].

Various molecular modelling programs are available which use different algorithms such as distance geometry (DG) and restrained molecular dynamics (rMD), to perform structure calculations. DG calculations are advantageous for structures of small molecules due to the speed at which structures can be generated (hundreds of structures can be generated in a few hours)[128], however DG algorithms are purely mathematical procedures, which generally only sample a very limited region of conformational space[129]. This can result in energetically incorrect, extended structures being produced[129]. rMD approaches are real space methods and generally produce more energetically correct structures as there is a better sampling of conformational space associated with these methods.

3.3.1 Simulated annealing

Simulated annealing is a widely used rMD technique, particularly for nucleic acid structures as it allows the generation of multiconformer models[130]. During simulated annealing calculations the system is elevated to a high temperature which provides the system with increased kinetic energy, allowing a wider conformational space to be explored and local energy barriers to be overcome[130]. The system is then gradually cooled in order to trap the system in the lowest energy conformation[126, 131].

A potential problem with simulated annealing is that the structure can become 'stuck' in local energy minima (Figure 3.3); to overcome this a target temperature must be chosen which is high enough to overcome smaller energy barriers but low enough to ensure that once the system has reached the global minimum it cannot 'climb out'[130].
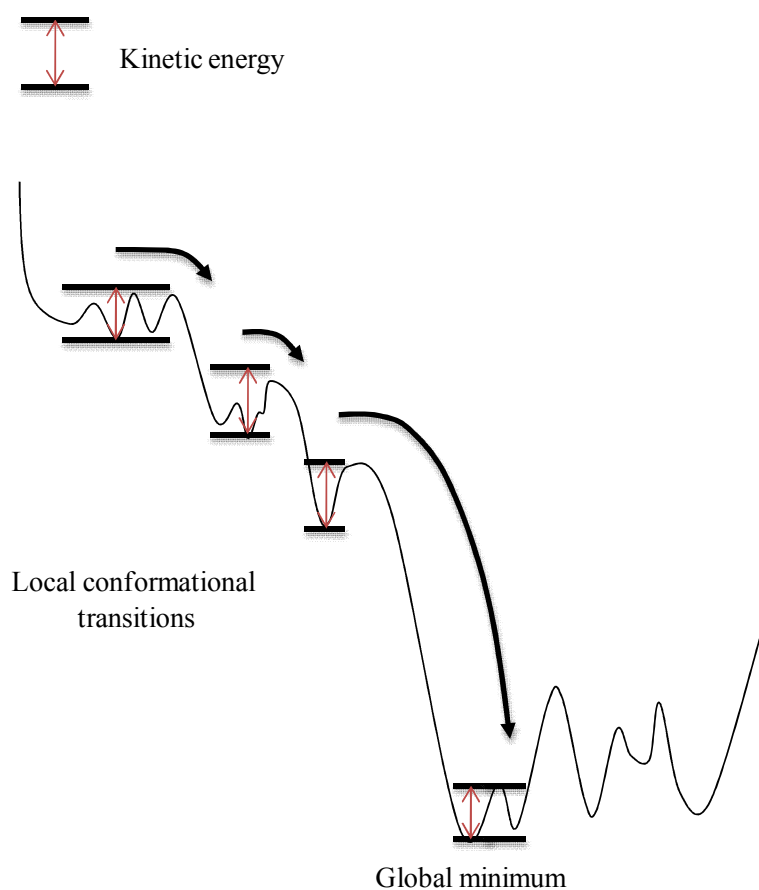
Figure 3.3. A schematic representation of the simulated annealing process.

Simulated annealing uses both experimental constraints and non-experimental information in the form of the force field, to search all the conformational space available and define all structures which satisfy the constraints. The force field provides necessary information regarding the stereochemistry of the molecule, such as bond lengths, angles, planarity and chirality[127] and is described in Equation 3.3.

$$V_{total} = V_{bond} + V_{angle} + V_{dihedral} + V_{VdW} + V_{coulomb} + V_{H\ bond} + V_{nOe} + V_{J\ coupling}$$

Equation 3.3

The first five terms monitor the classical potential energy of the molecule. $V_{H\,bond}$ is used to maintain hydrogen bonds within the structure and may be experimentally justified (for example, nOe distances obtained from NOESY cross peaks corresponding to imino protons) or known values from x-ray crystallography studies. The final two terms are energy penalties which are applied to conformations which do not fulfil the experimental restraints[119].

To perform simulated annealing calculations, a starting structure is required which is usually the extended or unfolded structure. From this, a converged structure is produced which is then energy minimised to give a final structure (Figure 3.4). Initially, only nOe and hydrogen bonding constraints are used to maximize the search of all available conformational space, and then additional MD calculations are performed which include dihedral angle constraints. Van der Waals and electrostatic energy terms are then added during the final energy minimisation step[111].
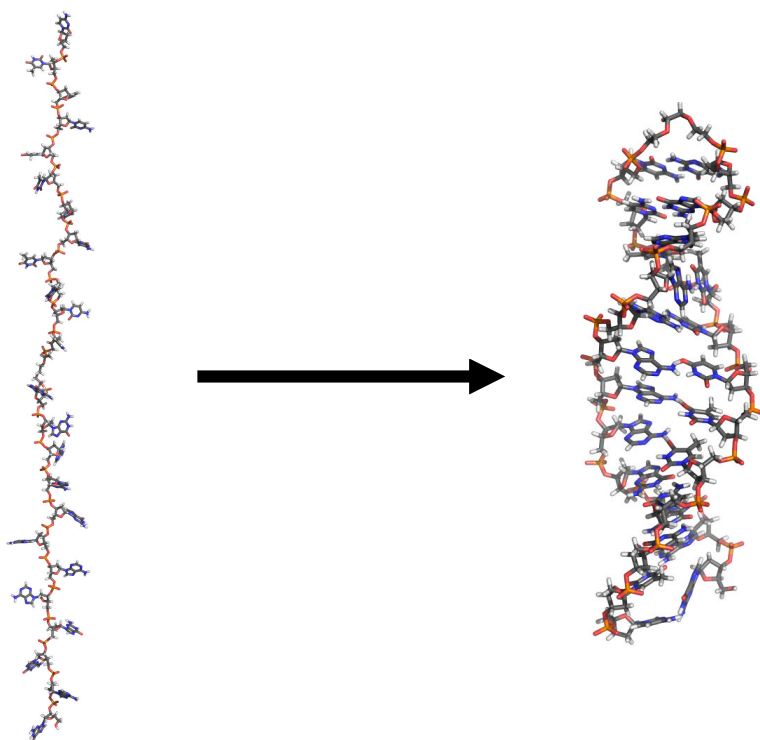
Figure 3.4. Illustration of simulated annealing of the hairpin duplex, starting from the extended structure (left) from which an energy minimised, helical structure was produced (right).

3.4 CNS-solve

There are various molecular modelling programs available for structure generation which differ in the algorithms which are used to perform the MD calculations. CNS-solve, the successor to X-PLOR[132] is well adapted for use in the structure determination of nucleic acids, and can incorporate both crystallographic and NMR-based constraints. For NMR constraints this may include nOe distances, ^{1}H and ^{13}C chemical shifts, J-coupling data, hydrogen bond distances, dihedral angles and residual dipolar couplings.

3.4.1 Annealing protocol

The simulated annealing protocol implemented in CNS-solve consists of four phases. The first is a high temperature phase using either torsion-angle or Cartesian molecular dynamics. The structures are then subjected to a first torsion-angle or Cartesian slow cooling stage, followed by a second, Cartesian slow cooling phase. Finally, the structures are subjected to a conjugate-gradient minimisation step to produce the final, energy minimised structures.[133]

The NMR structure calculation is performed as a hybrid-energy-function-based optimisation problem using the following energy function (Equations 3.4 - 3.6).

$$E = E_{chem} + E_{NMR}$$ **Equation 3.4**

$$E_{NMR} = \omega_{nOe} E_{nOe} + \omega_{dihedral} E_{dihedral}$$ **Equation 3.5**

$$E_{chem} = E_{geom} + \omega_{vdW} E_{vdW}$$ **Equation 3.6**

E_{chem} describes the agreement between actual and expected values for bond lengths and angles, planarity, chirality and non-bonded interactions (van der Waals and hydrogen bonding interactions). E_{nOe} represents the nOe derived distance constraints, which are described by a flat-bottomed parabolic function (Equations 3.7 and 3.8)[133]. ω_{noe}, $\omega_{dihedral}$ and ω_{vdW} represent the weights for each of the energy terms.

$$E_{nOe} = min \begin{cases} \Delta^2 & d_{upper} + 0.5 > R \\ a + \dfrac{b}{\Delta} + \Delta d_{upper} & + 0.5 < R \end{cases}$$

Equation 3.7

$$\Delta = \begin{cases} (R - d_{upper}) & d_{upper} < R \\ 0 & d_{lower} < R < d_{upper} \\ (d_{lower} - R) & R < d_{lower} \end{cases}$$

Equation 3.8

R is the distance between a pair of protons, d_{upper} and d_{lower} are the upper and lower boundaries for R, and *a* and *b* are determined such that E_{nOe} is a differentiable function at the point R = d_{upper} + 0.5. The sum is carried out over all the nOes[133].

3.5 Structure calculation for the hairpin duplex

Two sets of structures were generated for the hairpin using both non-exchangeable and exchangeable proton distance restraints. Two different methods were employed for calculating non-exchangeable distances, first using a single NOESY spectrum recorded with a 200 ms mixing time, and then a range of spectra collected over a range of mixing times. One set of exchangeable proton distance restraints was used in both sets of structure calculations, using data taken from a single NOESY spectrum

recorded with a 250 ms mixing time. The quality of the structures were analysed and compared in order to determine the most appropriate method, which will later be used in the structure generation for the non-modified and modified triplexes (see Chapter 5).

3.5.1 Calculation of inter-proton distances

Non-exchangeable inter-proton distances were calculated from NOESY spectra recorded with mixing times of 50, 100, 150, 200, 250, 300, 350 and 400 ms for the hairpin duplex in D_2O. Cross peak volumes were obtained by integrating each peak, and build-up curves were constructed for each proton pair. The distance dependence was found to deviate from linearity at a mixing time of 250 ms, therefore it was decided that the 200 ms data set would be used to generate structures based on the ISPA model. The cytidine H5-H6 distance (2.4 Å) was chosen as the reference distance as it is covalently restrained and does not alter with conformation[110], and the volume of one of the C H5-H6 cross peaks was measured and used as the reference volume. C13 H5-H6 was used as the reference, as this cross peak was visible in every spectrum and there was no spectral overlap in this region. The build-up curve constructed for the reference cross peak is shown in Figure 3.5. By then measuring all of the assigned NOESY cross peaks in the spectrum and inputting the values into Equation 3.1, inter-proton distances could be calculated for many proton pairs.
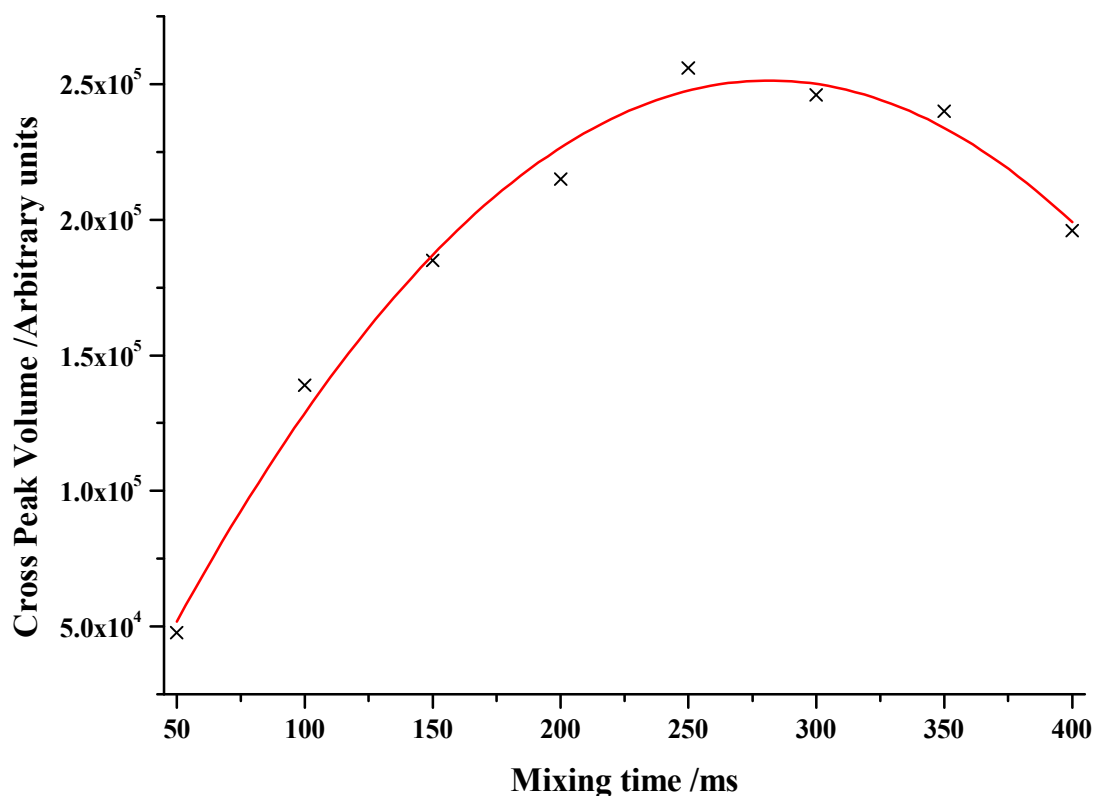
Figure 3.5. The build-up curve constructed for the hairpin C13 H5-H6 cross peak. The 200 ms peak volume was used as the reference volume.

The second set of structures were generated using nOe based distances that had been calculated from the gradient of the early, linear region of the build-up curves, which is equivalent to the cross relaxation rate. These rates were compared to the cross relaxation rate of the reference proton pair (C13 H5-H6, 2.4 Å) and distances calculated using Equation 3.1, where η_{ref} and η_{ab} correspond to the cross relaxation rates of the reference and unknown distances respectively.

Distances between proton pairs, in which one or both of the pair is a labile proton, were calculated from a NOESY spectrum (250 ms mixing time) of the hairpin acquired at low temperature, in a 9:1 H_2O:D_2O solvent mix. This set of nOes was

used for both structure generation attempts i.e. one based on nOe build-up, one based on nOe information taken from the 200 ms data set.

Loose errors of ± 30 % were applied to the non-exchangeable nOe distances for all residues excluding the terminal base pairs which were given boundaries of ± 50 % to account for end fraying. Error boundaries of ± 40 % were applied to exchangeable nOe distances in accordance with the literature[134]. The nOe restraints files used for the hairpin 200 ms and build up structures are provided in Appendix A1.

3.5.2 Additional restraints data

In addition to the nOe derived distances, hydrogen bonding and planarity constraints were specified in order to maintain the correct Watson-Crick base pairing. All hydrogen bonding distances were taken from standard values, with error boundaries of ± 35 % and included both the hydrogen bond donor and acceptor atoms[1]. The distances between the heavy atoms participating in the hydrogen bond were also restrained to ensure linearity of the bonds. Planarity restraints were applied to all except the terminal base pairs in order to keep the base pairs in a plane[135].

Dihedral angles restraints were also included in the structure calculations. Backbone torsion angles were loosely restrained to a range which encompasses typical values for both A- and B-type helices, in accordance with the literature[135, 136].

The experimental restraints used in the structure generation of the hairpin are detailed in Table 3.1. Although the average numbers of nOe restraints for both sets of structures were slightly less than the preferable value (13 were included for the 200

ms data and 12 for the build-up data), every residue was well defined by both inter-
and intra-residue distance constraints.

Table 3.1. Restraints used in the structure generation of the hairpin.

Constraint type	200 ms data set	Build-up data set
Non-exchangeable nOe	301	267
Exchangeable nOe	22	22
Hydrogen bond	60	60
Planarity	10	10
Torsion angle	125	125

3.5.3 Simulated annealing

Simulated annealing of the hairpin was initially performed from an extended starting structure using torsion-angle dynamics followed by Cartesian molecular dynamics. Parameters are detailed in Table 3.2. For each of the two data sets (200 ms and build up data) 10 rounds of annealing were carried out, each with a different initial velocity at which the two strands of the hairpin approached one another[137], and each round was executed until 100 'acceptable' structures were produced[138]. Acceptable structures are defined as those which contain no nOe violations greater than 0.5 Å, no dihedral angle violations greater than 5 °, root mean square deviation (RMSD) of bonds from ideal values less than 0.02 Å and RMSD of angles not exceeding 2.0 °[133]. The lowest energy acceptable structure was then taken as the new starting point for a further round of simulated annealing using Cartesian molecular dynamics[139]. An ensemble of the 10 structures with lowest total energy was then produced.

Table 3.2. Parameters used for each stage of initial simulated annealing of the hairpin structure.

	Stage 1	Stage 2	Stage 3	Stage 4
	High-temperature torsion-angle molecular dynamics	Slow-cooling torsion-angle molecular dynamics	Slow-cooling Cartesian molecular dynamics	1000 steps conjugate gradient minimisation
Temperature (K)	20,000	20,000→1,000	1,000→300	-
Time step	0.015	0.015	0.003	-
Δt	60	60	6	-
ω_{nOe}	150	150	150	150
$\omega_{dihedral}$	5	5	200	400
ω_{vdw}	0.1	0.1→1.0	1.0→4.0	4

3.6 Generated structures

Average structures for both the 200 ms and build-up nOe data were calculated from the 10 lowest energy structures that were produced. These are illustrated in Figure 3.6 and Figure 3.7 respectively. Table 3.3 shows the RMSD from ideal values for the average lowest energy structures. The average structures both have low RMSD values which are significantly below the limits for acceptance and no nOe violations greater than 0.5 Å or dihedral violations greater than 5 °. Both structures adopt a right handed helical structure as would be expected for an unmodified DNA duplex.

Table 3.3. RMSD from ideal values for the average of the ten lowest energy 200 ms and build-up structures.

Average Structure	RSMD of bonds / Å	RMSD of angles / °
200 ms	0.0023	0.4612
Build-up	0.0025	0.4850
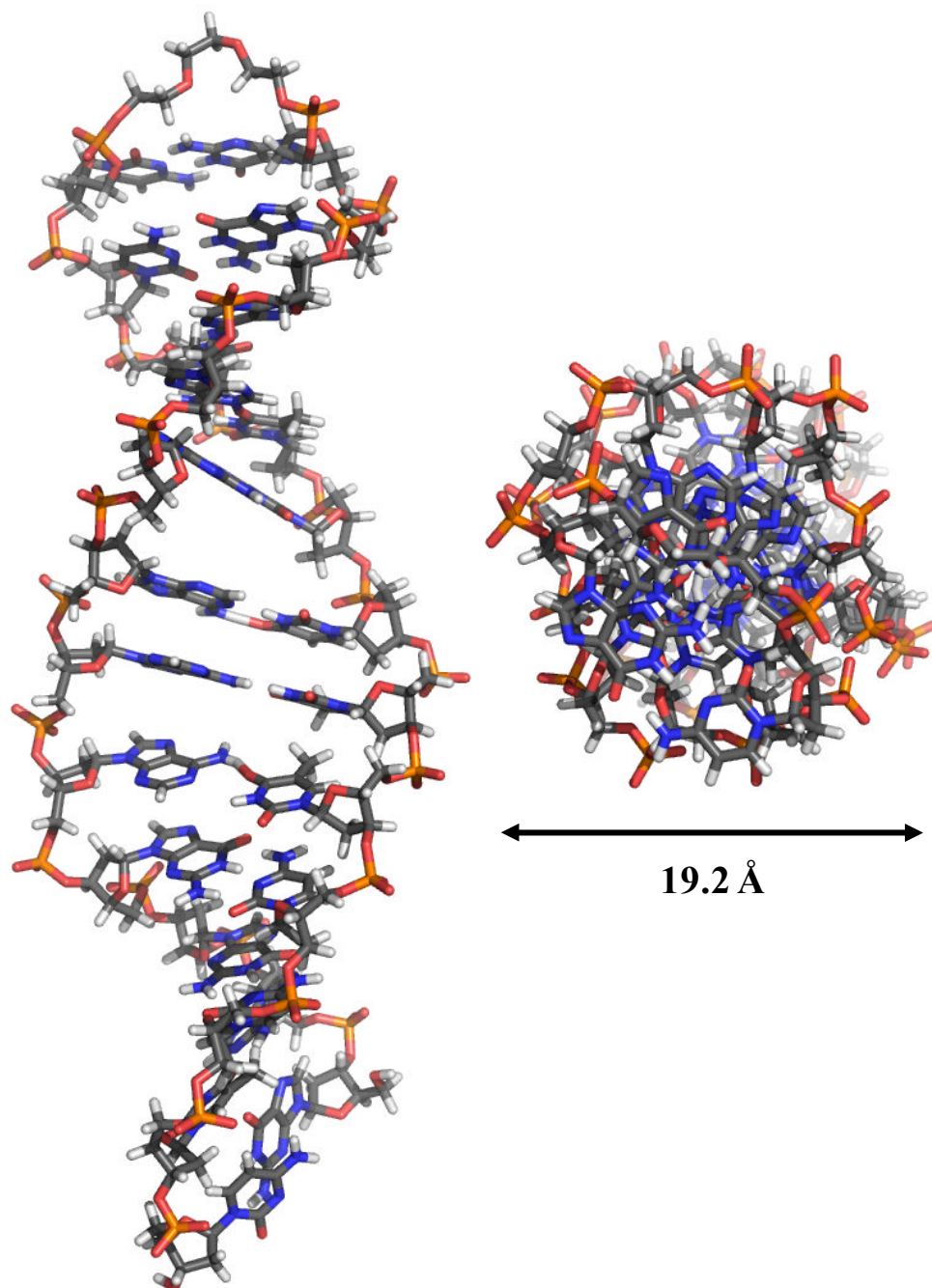
Figure 3.6. Average of the 10 lowest energy structures calculated from the 200 ms nOe data. View from the side is shown on the left and view from along the helical axis on the right. The helix diameter was measured at 19.2 Å.
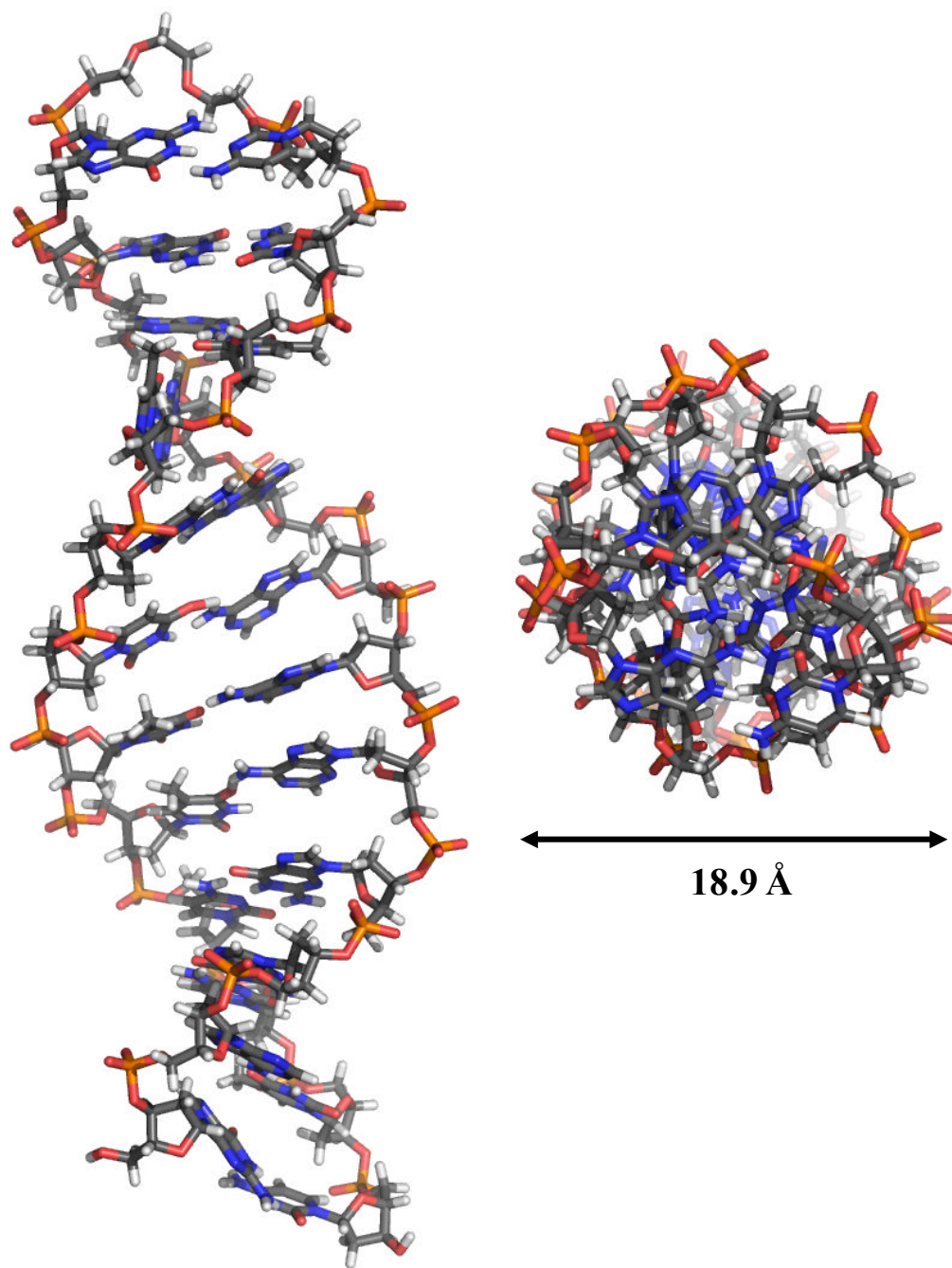
Figure 3.7. Average of the 10 lowest energy structures calculated from the build up nOe data. View from the side is shown on the left and view from along the helical axis on the right. The helix diameter was measured at 18.9 Å.

3.7 Evaluation of structures

3.7.1 Energy terms

The energy terms provide a measure of the quality of the calculated structures, with lower energies being more favourable. The NMR constraint energy (E_{NMR}), comprising the nOe energy (E_{nOe}) and dihedral angle energy (E_{dihed}) terms (Equation 3.9) describe how well the structure fits the experimental data; lower values indicate a better fit.

$$E_{NMR} = E_{nOe} + E_{dihed} \qquad \textbf{Equation 3.9}$$

The geometric energy (E_{geom}) describes how strained the geometry of the structure is (Equation 3.10). A large geometric energy value may signify the presence of a novel structure or incorrect conformation.

$$E_{geom} = E_{bond} + E_{angle} + E_{improper} + E_{planarity} \qquad \textbf{Equation 3.10}$$

The non-bonding energy term ($E_{non\text{-}bond}$) is a sum of the van der Waals (E_{VdW}) and electrostatic ($E_{electrostatic}$) energy terms (Equation 3.11). Favourable values are indicative of a deep energy minimum, and are usually observed when a large number of experimental constraints are used in the structure calculations. However, because solvent is neglected in these calculations, electrostatic interactions are excluded[133].

$$E_{non-bond} = E_{vdW} + E_{electrostatic}$$ **Equation 3.11**

The sum of all the above energy terms (Equation 3.12) gives the total energy (E_{total}), which again is more favourable for structures determined from extensive experimental restraints.

$$E_{total} = E_{NMR} + E_{geom} + E_{non-bond}$$ **Equation 3.12**

The energy values for the hairpin 200 ms and build-up structures are given in Table 3.4. For comparison, energy values for a B-type 14mer DNA duplex that had been calculated using CNS-solve are also provided[140].

Table 3.4. Energy values (in kJ mol^{-1}) for the average lowest energy 200 ms and build-up structures. Values for a B-type DNA duplex[140] are also given for reference.

Energy	200 ms structure	Build-up structure	B-type duplex[140]
Total	181.4	199.3	899.0
nOe	20.9	17.8	35.0
Dihedral	3.3	6.7	1.1
Bond	9.0	11.3	16.7
Angle	60.6	68.0	1751.0
Improper	9.3	10.9	19.4
Planarity	0.8	1.3	-
Van der Waals	77.5	83.4	− 434.0

Both hairpin structures have favourable total and non-bonding energies, suggesting that the structures were determined from an extensive set of NMR constraints. The reference data set has a much higher total energy, which implies that the nOe restraints were less extensive or of poorer quality. The B-type 14-mer duplex structures were calculated from fewer nOe restraints than both sets of hairpin structures (10 compared to 13 and 12 restraints per residue for the hairpin structures).

The two sets of hairpin structures have very similar energy terms, indicating that both the 200 ms and build-up methods for calculating nOe distance restraints are valid.

3.7.2 Atomic RMSDs

Successful refinement should produce similar final structures from different starting structures i.e. with small atomic RMSD values between the individual structures; values < 1.0 Å are generally considered adequate[119]. The ten lowest energy structures for the 200 ms and build-up data were superimposed (Figures 3.8 and 3.9), and atomic RMSD differences were calculated for the ensembles (Table 3.5).

Both sets of structures have slightly larger RMSD values at the ends of the molecule which is to be expected, since looser constraints were applied to the terminal residues to account for fraying at the ends of the molecule. Overall, both sets of structures have very low RMSD values, signifying that the structures generated have excellent convergence. The 200 ms structures have slightly lower values than the build up structures, which suggests that the cross relaxation rate method for analysing nOe constraints is less reliable than the method based on the ISPA model. This could be due to slight errors in cross peak measurement at short mixing times due to inherent noise[111].
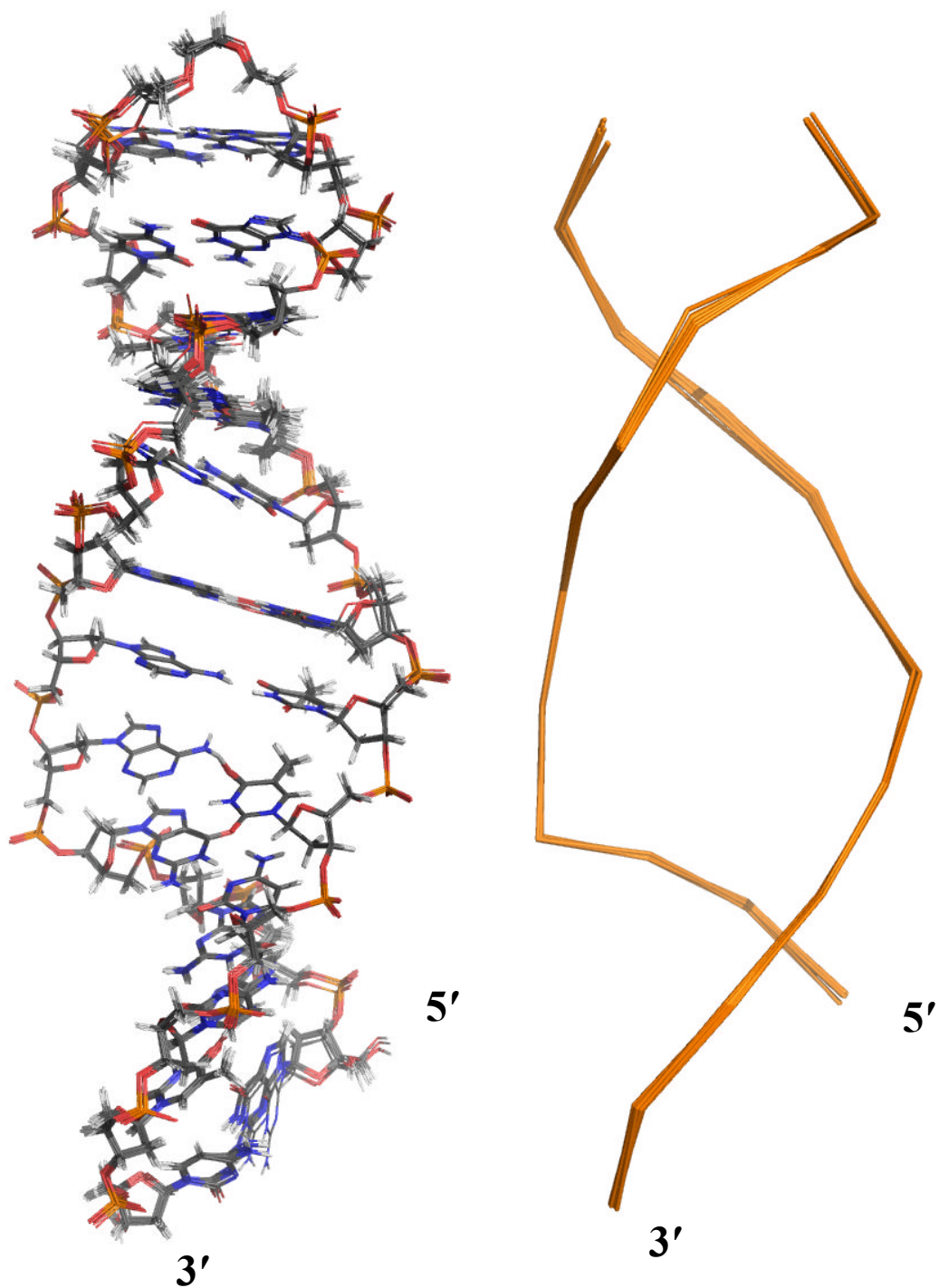
Figure 3.8. Superposition of the 10 lowest energy structures generated from the 200 ms data set. The molecular structure is shown on the left and a ribbon representation of the sugar phosphate backbone on the right.
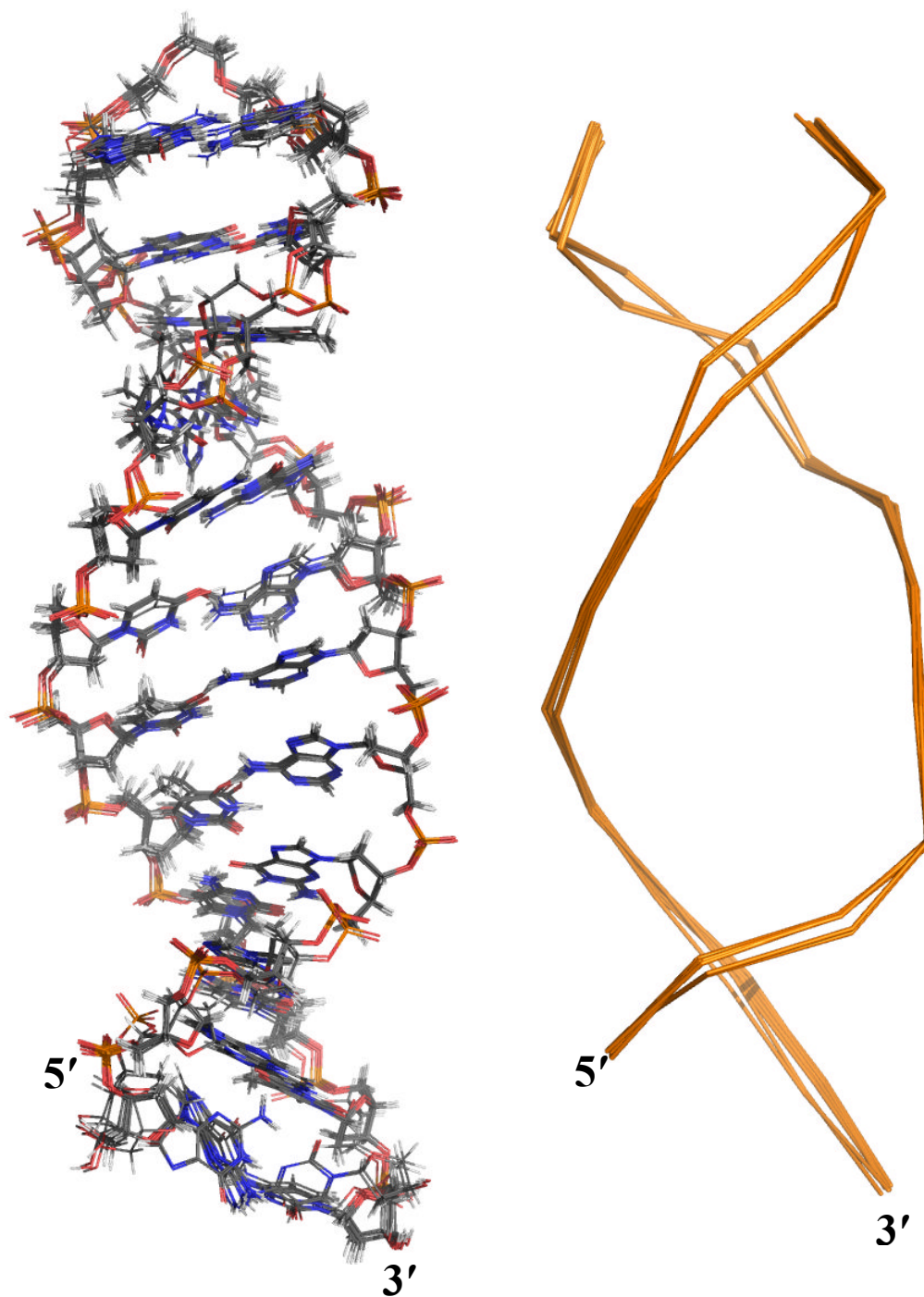
Figure 3.9. Superposition of the 10 lowest energy structures generated from the build-up data set. The molecular structure is shown on the left and a ribbon representation of the sugar phosphate backbone on the right.

Table 3.5. Atomic RMSD values for the ten lowest energy hairpin structures, on a per residue basis.

Residue	200 ms structures	Build-up structures
G1	0.428	1.212
A2	0.238	0.854
G3	0.125	0.290
G4	0.121	0.243
A5	0.140	0.230
A6	0.106	0.354
A7	0.156	0.351
G8	0.149	0.281
A9	0.242	0.994
A10	0.345	0.718
G11	0.137	0.656
G12	0.232	0.362
C13	0.485	0.407
C14	0.221	0.428
T15	0.349	0.837
T16	0.231	0.880
C17	0.155	0.259
dU18	0.315	0.224
T19	0.223	0.287
T20	0.150	0.346
C21	0.125	0.260
C22	0.124	0.300
T23	0.173	0.319
C24	0.225	0.966
Average	**0.216**	**0.493**

3.7.3 Torsion angle analysis

The backbone torsion angles for the average structures were obtained using the program CURVES[+141] and are given as an average for all residues in Table 3.6. The expected values for standard A- and B-type helices are also given. Figure 3.10 illustrates the torsion angles which describe the phosphodiester backbone in an oligonucleotide.

Table 3.6. Average backbone torsion angles (in degrees) for the lowest energy hairpin structures. Typical values for standard B-type[142] and A-type[143]duplexes are also given.

Structure	α	β	γ	δ	ϵ	ζ	χ
Hairpin 200 ms	− 49	− 138	40	141	− 140	− 95	− 69
Hairpin build-up	− 62	− 147	38	138	− 143	− 95	− 86
B-type DNA	− 46	− 147	36	157	− 133	− 96	− 98
A-type DNA	− 75	− 179	59	79	− 155	− 67	− 158
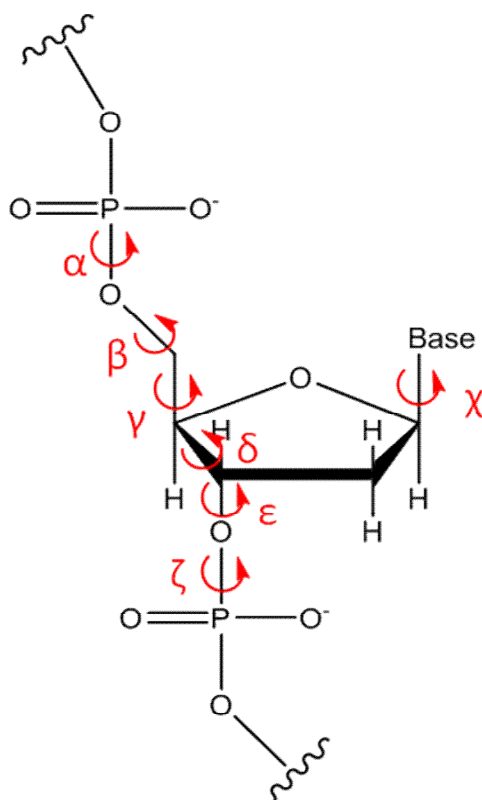
Figure 3.10. Illustration of the backbone torsion angles of an oligonucleotide chain.

The majority of torsion angles are comparable for both sets of structures, indicating that similar structures have been generated from the two data sets. On comparison to standard A- and B-type values, the torsion angles for both hairpin structures are more consistent with a B-type helix which is in agreement from the qualitative analysis of the NMR data detailed in Chapter 2. Deviations from ideal values may be due to differences in the nucleotide sequence.

3.7.4 Helical parameter analysis

The helix diameters were measured for the average 200 ms and build up hairpin structures. The helices were found to have an average diameter of 19.2, and 18.9 Å respectively, as indicated in Figure 3.6 and Figure 3.7. These values are much closer to the average helix diameter of 20 Å for B-type DNA than the slightly larger value of 23 Å for A-type helices[144]. The major groove widths were also measured for both sets of structures and are given in Table 3.7. The major groove widths are also indicative of a B-type helical structure.

Table 3.7. Average major groove widths for each set of hairpin structures. Values were calculated as refined P-P distances, taking into account the directions of the sugar-phosphate backbones[145]. Typical values for B-DNA[146] and A-DNA[147] are given for reference.

Structure	Major groove width /Å
Build-up	17.8
200 ms	16.8
B-DNA	17.3
A-DNA	13.6

Local base pair step parameters were determined from the CURVES$^+$ analysis of the average structures; average values for all base pair steps are given in Table 3.8 and the variation along the sequence is highlighted in Figures 3.11 and 3.12. For each parameter, values for the 200 ms structure are shown in red, the build up parameters are shown in blue and idealised values for a B-type DNA dodecamer[146] are given in green for comparison. A pictorial representation of each parameter is also shown in Figures 3.11 and 3.12.

From Table 3.8 it can be seen that the 200 ms structures have values closer to a standard B-type helix than the build up structures for all parameters except the roll and slide. For the 200 ms structures, average rise and twist parameters of 3.94 Å and 35.33 ° respectively are particularly close to ideal values, which are 3.36 Å (rise) and 35.58 ° (twist). Both the build up and 200 ms structures have a higher base pair tilt than a standard B-type helix and both also a greater base pair slide and shift. The build up structure has a negative average base pair roll of -3.46 °, whilst the 200 ms structures have a positive roll of 8.10 °, neither of which is comparable to the expected base pair roll for a B-helix of 0.02.

Table 3.8. Average base pair step parameters for the lowest energy build up and 200 ms structures. Values for an idealised B-type helix are also given for comparison[146].

Base pair step parameter	Build up structures	200 ms structures	Ideal B-type helix
Tilt (°)	9.83	4.71	− 0.19
Roll (°)	− 3.46	8.10	0.02
Rise (Å)	4.22	3.94	3.36
Twist (°)	24.38	35.33	35.58
Slide (Å)	0.40	0.86	0.12
Shift (Å)	1.26	0.92	− 0.02

Base Pair Tilt (in Degrees)

Base Pair Roll (in Degrees)

Base Pair Rise (in Angstroms)
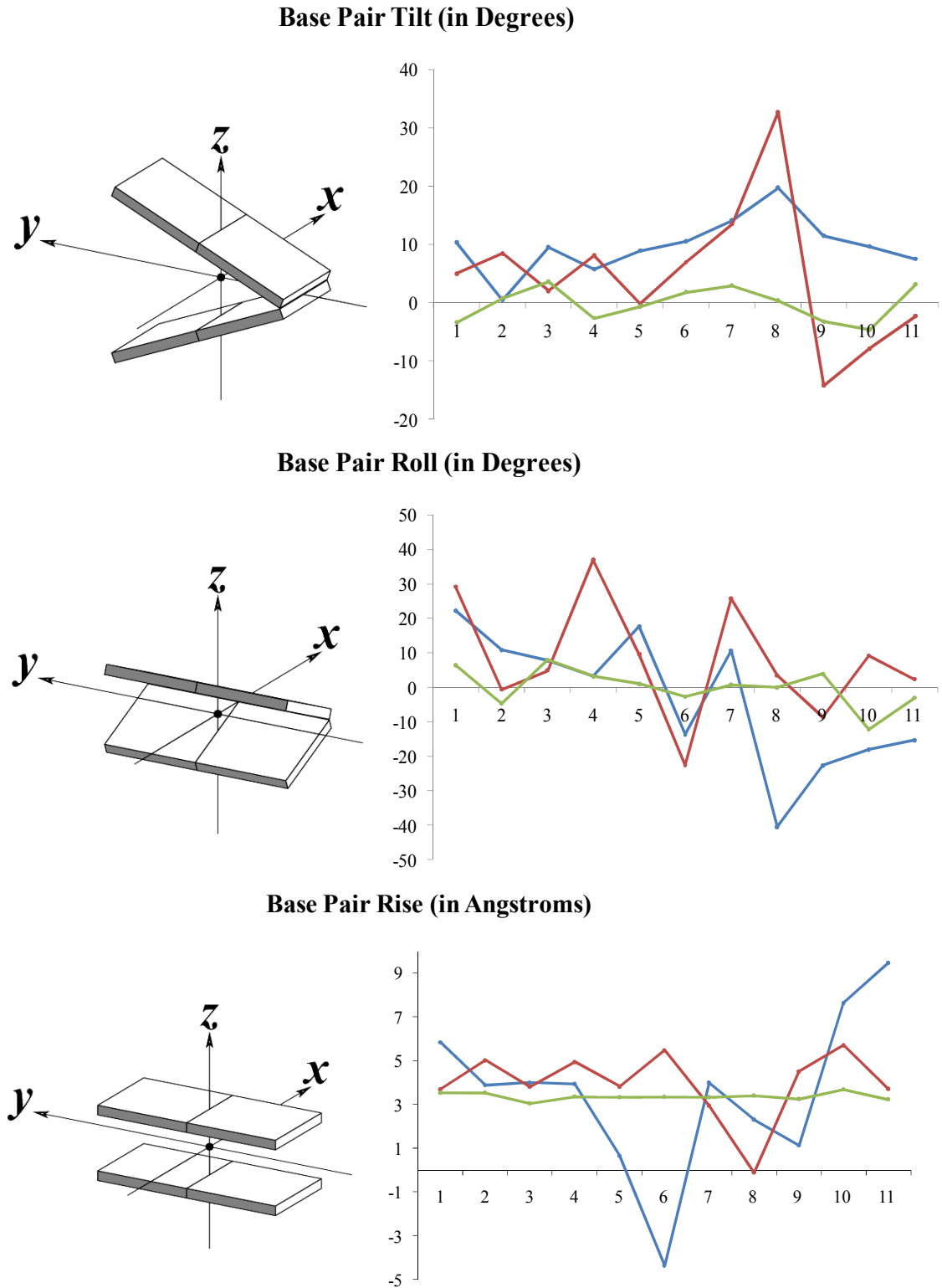
Figure 3.11. Base pair tilt, roll and rise values for the minimum energy hairpin 200 ms structure (blue), build-up structure (red) and an ideal B-type helix[146] (green). Numbers along the x-axis indicate base pair steps. Representations of positive values of each designated parameter[148] are also illustrated on the left.

Base Pair Twist (in Degrees)

Base Pair Slide (in Angstroms)

Base Pair Shift (in Angstroms)
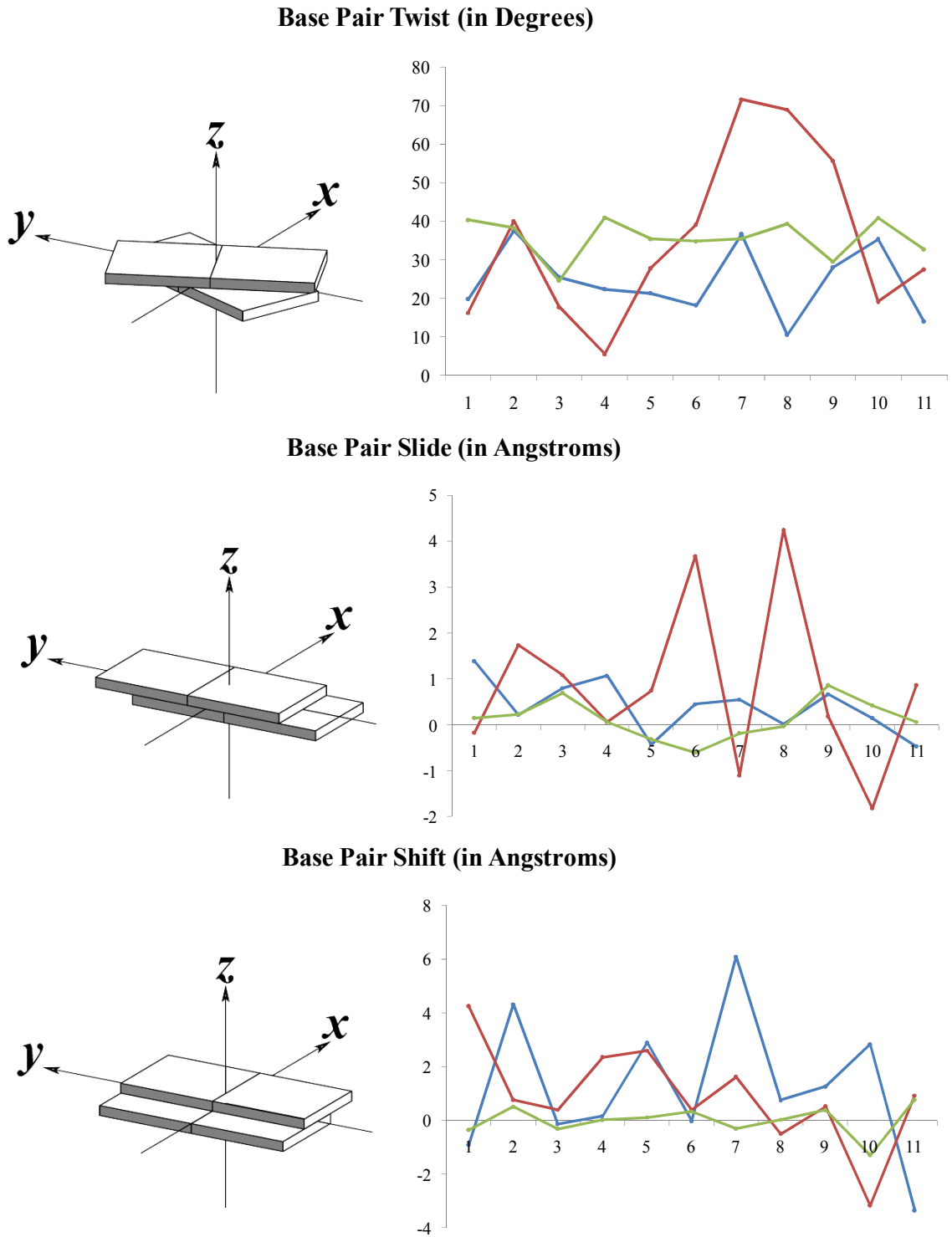
Figure 3.12. Base pair twist, slide and shift values for the minimum energy hairpin 200 ms structure (blue), build-up structure (red) and an ideal B-type helix[145] (green). Numbers along the x-axis indicate base pair steps. Representations of positive values of each designated parameter[148] are also illustrated on the left.

From Figures 3.11 and 3.12 it is clear that whilst most average base pair step parameters are comparable to a standard B-type helix, the local structure varies significantly along the oligonucleotide for both the 200 ms and build-up structures, with the latter showing a greater variation along the sequence. However, both structures do generally show the same trends and variations, signifying a good convergence between the two. Greater variation in all base pair step parameters is observed at both ends of the molecules which may be a result of looser constraints which were applied at the ends of the molecules to account for end fraying.

Deviations from ideal values are apparent for both structures which may be attributed to differences in the base pair composition of the hairpin sequence compared to the reference B-type helix; this is expected as helical parameters are sequence dependent[9].

From each graph it is apparent that the 200 ms structures have individual base pair step parameters that most closely resemble those for a B-type helix than the build up structures, which deviate significantly. As the qualitative NMR analysis detailed in Chapter 2 concluded that the hairpin adopts a standard B-type structure, no significant deviations from ideal values other than those due to sequence variations are expected. This suggests that the accuracy of the lowest energy 200 ms structure is greater than the lowest energy build up structure, which may be a result of how well defined the structures were by nOe constraints - the 200 ms structure was better characterised by nOe constraints than the build up structure (323 200 ms exchangeable and non-exchangeable proton nOe constraints were used compared to 289 build up constraints).

It is also worth noting that the largest fluctuations in base pair step parameters generally occur around base pair 8. Only 6 (200 ms) and 7 (build up) non-exchangeable nOe constraints were employed for the cytosine residue in base pair 8 due to significant overlap between cross peaks in this region of the NOESY spectra. This prevented the integration of several cross peaks and so a number of proton distances could not be calculated for this nucleotide.

3.8 Discussion

The structures generated have favourable energies and no restraint violations, with backbone torsion angles consistent with expected values for B-type DNA. The pairwise RMSD values describe the precision of a set of structures, and the low values that were achieved for both the 200 ms and build-up structures indicate that the structures are precisely defined and only one conformation is present[111].

The number of restraints per residue was reported as a measure of the accuracy of the structures, and although the values of 13 (200 ms data) and 12 (build-up data) were slightly lower than ideal all residues appeared to be well defined, suggesting that the generated structures are a good representation of the true hairpin. However, this method has several limitations as restraints can be dependent on the size and shape of the molecule, and the residue-type composition[149], and proper assessment of the accuracy of molecular structures requires further validation of the experimental constraints.

Although many structure validation tools are available for the analysis of protein structures, validation of nucleic acids structures is virtually non-existent due to the lack of reference data in nucleic acid data bases. Validation was therefore limited to

describing various structural features such as backbone torsion angles and helical parameters, and comparing with typical values for standard A- and B-type duplexes. Both the 200 ms and build-up structures had features which were more consistent with those for B-type helices, which is in agreement with the NMR analysis presented in Chapter 2. The major groove widths were also measured and found to be consistent with B-type structures.

Helical parameters for both structures were generally comparable to those for a standard B-type DNA dodecamer, however greater variations along the sequence were observed. The greatest variations occurred in the section of the molecule which was less well defined by nOe constraints, indicating that the number of experimental constraints can and do affect the quality of a structure.

Structures calculated from both the 200 ms and build-up nOe restraints had comparable energy values and backbone torsion angles, indicating that use of the different restraints files produces similar final structures and therefore both methods for determining nOe distances are valid. Of the two, the 200 ms data produced structures with slightly better convergence and precision, as indicated by the pairwise RMSD values. Again this is likely to be a result of the greater number of nOe constraints used in generating the 200 ms structure set.

The nOe build-up method should theoretically produce better quality structures than the 200 ms method, as it takes into account all possible relaxation pathways and so provides a better reflection of the relaxation matrix intensities. However experimentally this has proved not to be the case, and although both methods produced similar and acceptable structures the 200 ms structures were of a slightly better quality. Although the 200 ms method which is based on the ISPA model, does

not take spin diffusion effects into account, spin diffusion can still be considered negligible at the relatively short mixing time of 200 ms. As this method allows a greater number of proton distances to be calculated and used as nOe constraints, which seems to be more important for improving structure quality than the actual method used for calculating constraints, the 200 ms method will be used in further structure generation of the non-modified and modified triplex (see Chapter 5).

3.9 Conclusions

Low energy structures were generated for the hairpin using nOe distance constraints that were calculated by two different methods. Both methods produced similar structures that were of a high quality. All structures produced had favourable energies, no violating constraints and low RMSD values, with backbone torsion angles that were consistent with B-type DNA.

Simulated annealing of the hairpin using both the 200 ms and build-up nOe data consistently produced similar structures, as can be seen from the superpositions of the lowest energy structures and the RMSD differences between heavy atoms. It can therefore be concluded that the simulated annealing protocol using torsion angle molecular dynamics was sufficient and that both methods for determining nOe restraints are acceptable.

Overall, the 200 ms data produced more consistent structures, and so this method will be used in performing structural analysis of the native and modified triplexes.

Chapter 4

NMR Analysis of the Native and Modified Triplexes

Chapter 4. NMR Analysis of the Native and Modified Triplexes

4.1 Introduction

With the NMR analysis and structure determination of the target hairpin complete (Chapters 2 and 3, respectively), it was possible to progress to analysis of both the native and modified triplexes. The triplex under investigation was described in Chapter 1 and the sequence and numbering is given in Scheme 4.1.

$$5'\text{-} C_{25}^+ \, T_{26} \, C_{27}^+ \, C_{28}^+ \, T_{29} \, T_{30} \, T_{31} \, C_{32}^+ \, T_{33} \, T_{34} \, C_{35}^+ \, C_{36}^+ \text{-} 3'$$

$$5'\text{-} G_1 \, A_2 \, G_3 \, G_4 \, A_5 \, A_6 \, A_7 \, G_8 \, A_9 \, A_{10} \, G_{11} \, G_{12}$$

$$3'\text{-} C_{24} \, T_{23} \, C_{22} \, C_{21} \, T_{20} \, T_{19} \, dU_{18} \, C_{17} \, T_{16} \, T_{15} \, C_{14} \, C_{13}$$
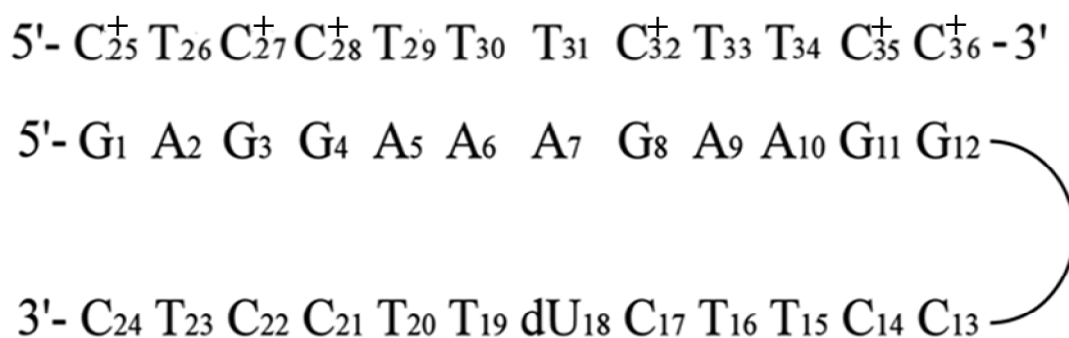
Scheme 4.1. The sequence and numbering of the triplex under investigation.

DNA triplexes, like DNA duplexes generally exhibit a right handed helix[150], therefore similar NMR resonance assignment procedures were followed and sequential through-space connectivities were observed. Non-exchangeable protons were assigned using a series of two dimensional NOESY, TOCSY and DQF-COSY spectra which were acquired for the triplex sample in a D_2O solution. Exchangeable proton resonance assignments were carried out by obtaining a NOESY spectrum for the triplex in a 9:1 H_2O:D_2O solvent mix. NMR experiments were performed at pH 5

in order to fully protonate the third strand cytidines thus enabling Hoogsteen bonding between that strand and the purine strand of the duplex.

4.2 Resonance assignments for the native triplex

Evidence for the formation of the triplex was provided by analysing the exchangeable proton spectrum of the molecule (Figure 4.1). Broad, downfield shifted resonances at 14 - 16 ppm are characteristic of imino resonances from protonated cytidine residues in the third strand and downfield shifted peaks in the region of 9.5 - 10.5 ppm correspond to C^+ amino protons[151].
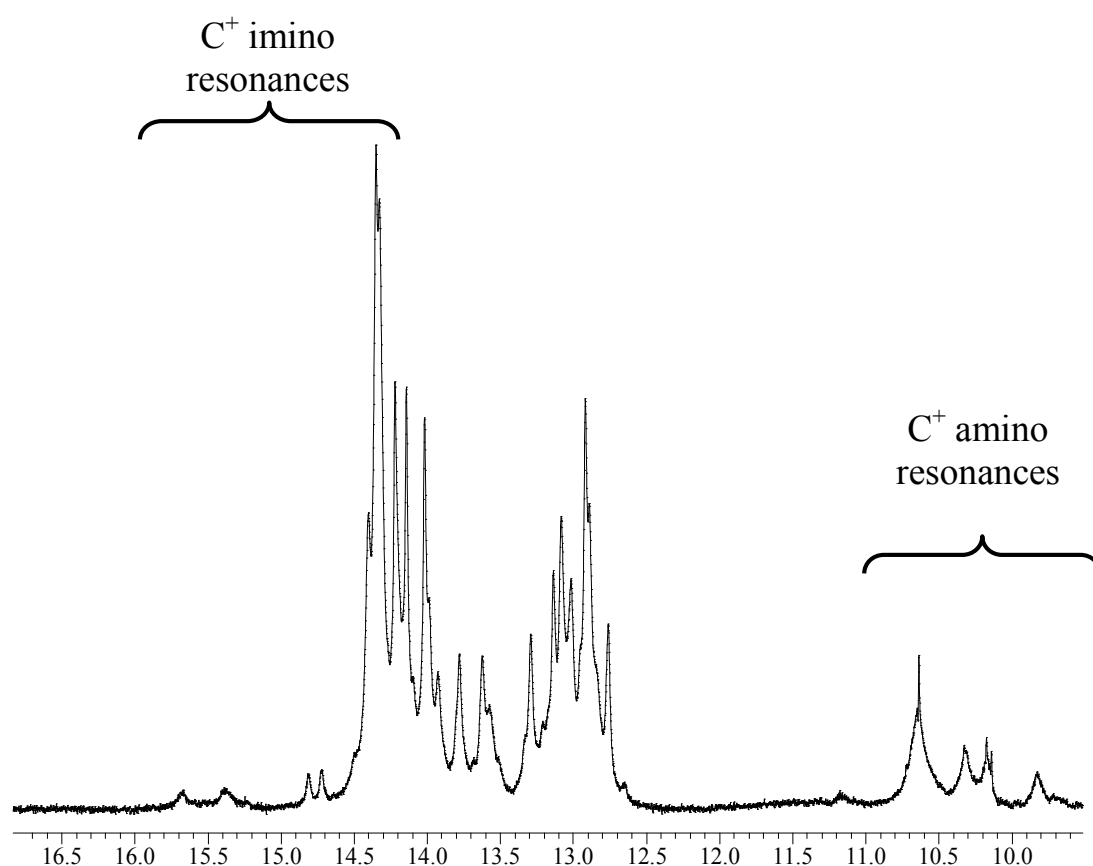
Figure 4.1. Expansion of the one dimensional ^{1}H spectrum of the non-modified triplex at pH 5, recorded in a 9:1 H$_2$O:D$_2$O solution mix at 10 °C. The lowfield imino and C^+ amino resonances that are characteristic of a parallel triplex are highlighted.

4.2.1 Assignment of exchangeable proton resonances

The exchangeable proton region of a 2D NOESY spectrum can provide a useful starting point for the triplex NMR assignments. Inter-triplet nOe connections (Figure 4.2) serve as evidence for binding of the third strand to the hairpin by Hoogsteen hydrogen bonds.
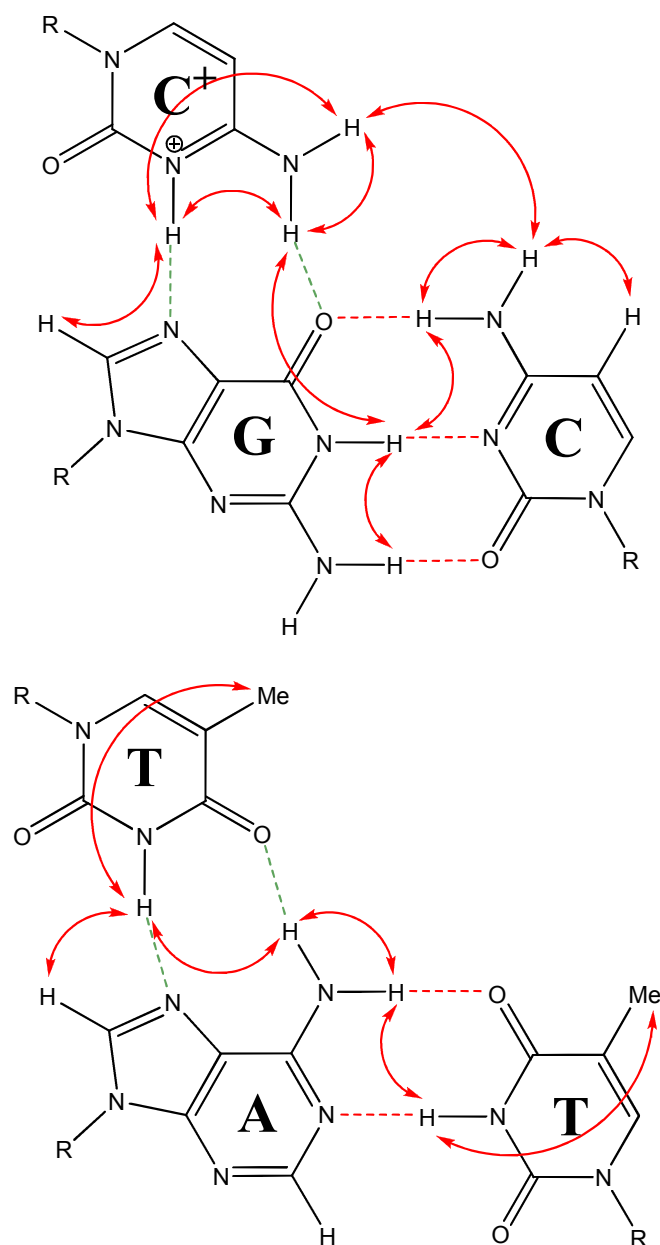
Figure 4.2. Expected nOe connections for exchangeable protons within C⁺:G-C and T:A-T base triplets.

Six C^+ imino resonances were expected, corresponding to the six protonated cytidine residues in the third strand. Of these, four could be identified on the NOESY spectrum; two C^+ imino protons by their connections to the intranucleotide pair of C^+ amino protons and another two by their lowfield chemical shifts at 14.6 and 14.7 ppm. Two signals which were apparent on the 1D spectrum at 15.4 and 15.7 ppm were very broad and connections to these could not be detected under any of the experimental conditions tried; this is consistent with terminal exchangeable protons. It was therefore assumed that these signals correspond to C26 and C34.

The two C^+ imino protons resonating at 14.6 and 14.7 ppm did not display any nOe connections to other imino protons, and so a complete set of sequential imino-imino connectivities between C^+ and T imino protons along the Hoogsteen strand could not be traced. C32, T31 and T33 imino protons could, however be assigned based on two sequential connections from a cytidine to two thymidine imino protons. C32 is the only residue within the third strand which has two neighbouring T residues, and so these could be unambiguously assigned. T30 could then also be identified via an nOe cross peak from its imino proton to the imino of T31 (Figure 4.3).
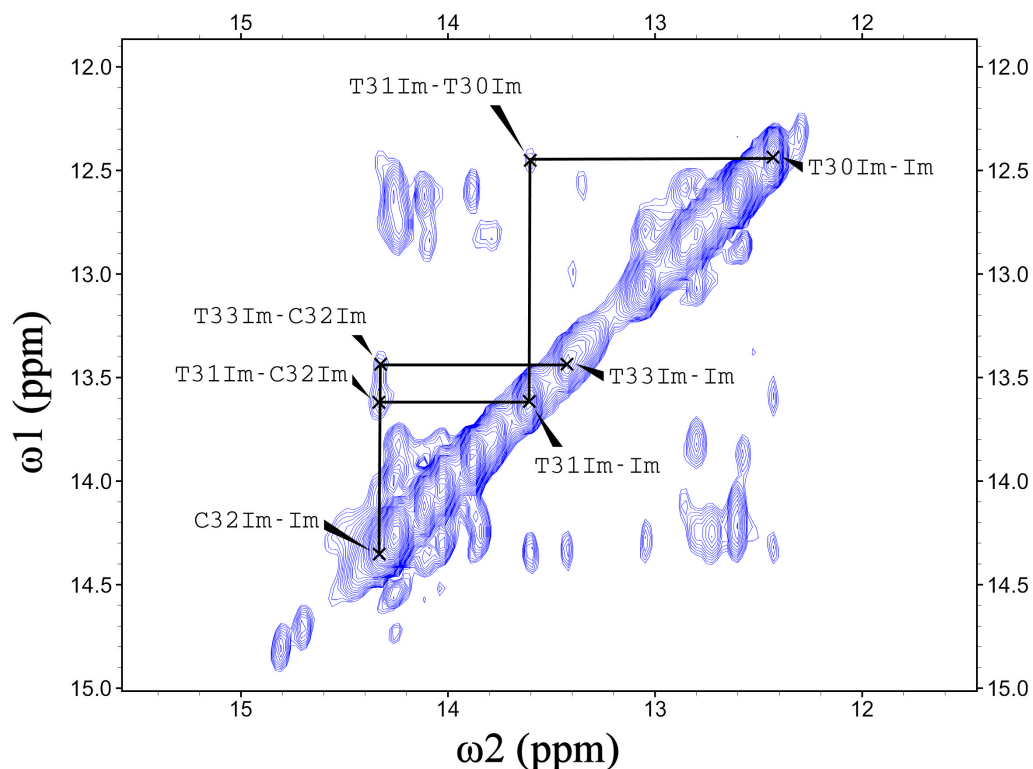
Figure 4.3. Expansion of the NOESY spectrum for the non-modified triplex in a (9:1) H$_2$O:D$_2$O solution, showing connections between T30, T31, C32 and T33 imino protons.

The remaining T imino protons from the Hoogsteen strand, which resonate at a similar chemical shift to the G imino protons (12 - 13.5 ppm) could be readily identified through nOe connections to their own methyl protons. Likewise, T imino protons from the Watson-Crick duplex exhibited similar connections to their methyls. A full set of sequential imino-imino connectivities could then be traced, and were attributed to the Watson-Crick base paired iminos. This facilitated the assignment of all residues within the hairpin with the exception of the terminal residues, G1 and G12 (Figure 4.4).
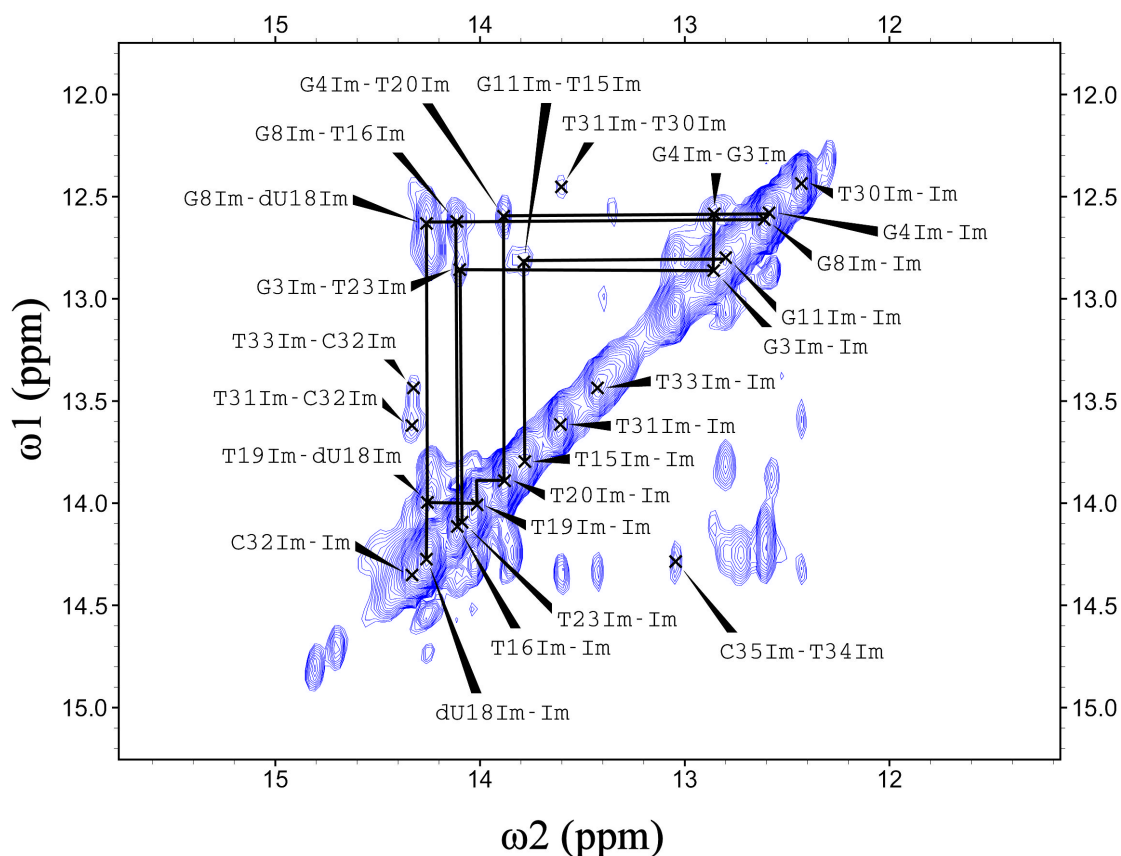
Figure 4.4. Expansion of the NOESY spectrum for the non-modified triplex in a (9:1) H₂O:D₂O solution, showing connections between the Watson-Crick related imino protons.

Several inter-strand nOe cross peaks connected imino protons from the purine strand of the duplex to unassigned Hoogsteen bonded iminos, facilitating the remaining imino assignments as well as providing further evidence that the third strand is interacting with the hairpin in a triple helical manner. The imino correlations observed for the native triplex are highlighted in Figure 4.5.
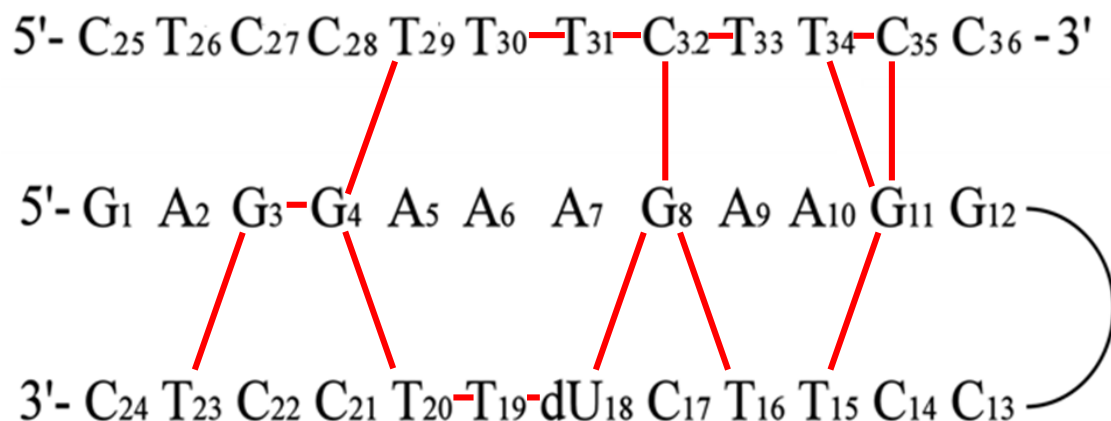
Figure 4.5. nOe connections between imino protons that were observed for the non-modified triplex structure.

Several cross peaks were also observed which connected pyrimidine imino protons from the Hoogsteen strand to the H2′ and H2″ protons on neighbouring purine residues in the 5′ direction. Such cross peaks are not observed in B-DNA duplexes but are characteristic of triplex formation[152]. These are shown in Figure 4.6 and provided a useful starting point for the assignment of the non-exchangeable protons (described in Section 4.2.2).
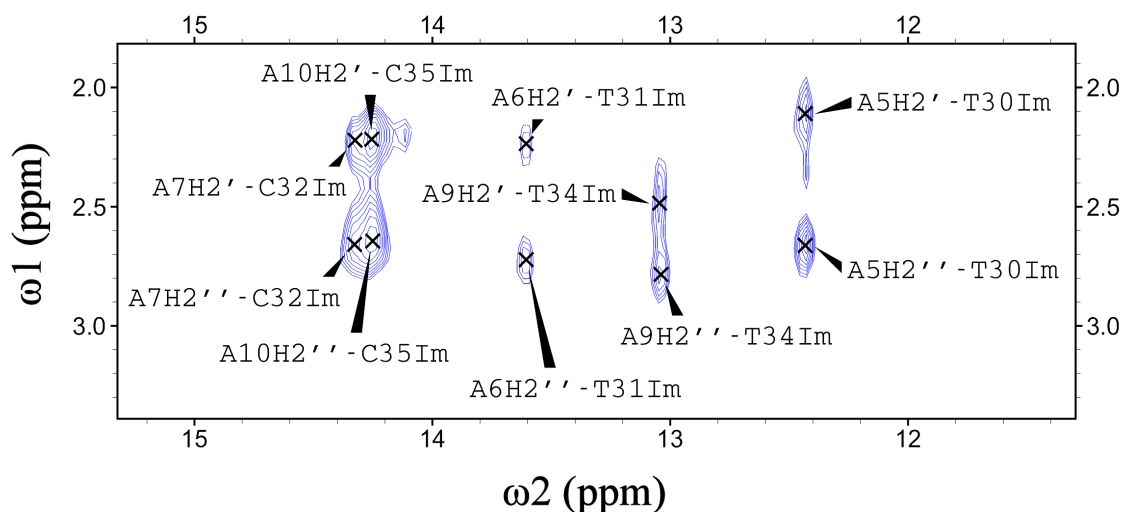
Figure 4.6. Expansion of the NOESY spectrum of the non-modified triplex in a (9:1) H$_2$O:D$_2$O solution, showing the region in which the third strand imino protons exhibit connections to 5′-neighbouring purine H2′/H2″ protons.

Assignment of some of the aromatic and amino resonances could then be obtained via intra-triplet nOe cross peaks to the assigned imino resonances[151, 153]; these are indicated in Figure 4.7. A5, A6, A7 and A9 H2 protons were identified via strong nOe connections to their Watson-Crick base paired T imino protons, and several A amino protons could also be assigned based on nOes to Hoogsteen bonded T imino resonances. Observed cross peaks between Watson-Crick paired G imino and C amino protons, and between third strand C$^+$ iminos and aminos completed the assignment of the exchangeable proton resonances for the non-modified triplex.

One final set of nOe cross peaks between imino protons and purine H8 resonances were identified and this provided an additional starting point for assignment of the non-exchangeable resonances (Section 4.2.2). The chemical shifts of the assigned labile and H2 protons are given in Table 4.1.
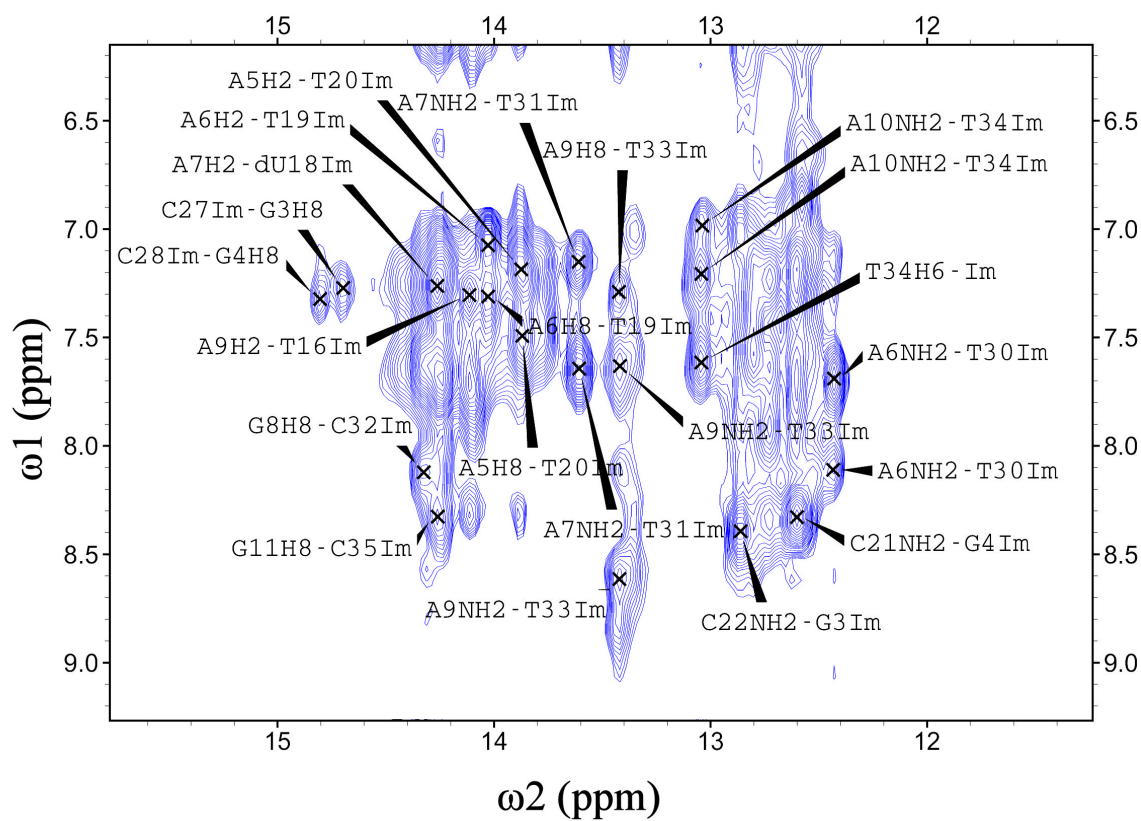
Figure 4.7. Expansion of the NOESY spectrum of the non-modified triplex in a (9:1) H₂O:D₂O solution, showing nOe connections between imino and amino/aromatic protons.

Table 4.1. Chemical shifts (in ppm) of all exchangeable and H2 protons for the native triplex at 2 °C, referenced to H₂O at 4.95 ppm.

Residue	Imino	Amino	H2
G1		-	-
A2	-	-	-
G3	12.86	-	-
G4	12.58	-	-
A5	-	-	7.12
A6	-	8.11 / 7.69	7.01
A7	-	7.64 / 7.29	7.27
G8	12.62	-	-
A9	-	8.63 / 7.63	7.29
A10	-	7.62 / 7.21	-
G11	12.80	-	-
G12	-	-	-
C13	-	-	-
C14	-	7.57 / 7.23	-
T15	13.80	-	-
T16	14.11	-	-
C17	-	-	-
dU18	14.27	-	-
T19	14.01	-	-
T20	13.89	-	-
C21	-	8.33	-
C22	-	8.39	-
T23	14.09	-	-
C24	-	-	-
C25	-	-	-
T26	12.74	-	-
C27	14.71	-	-
C28	14.80	-	-
T29	13.34	-	-
T30	12.44	-	-
T31	13.62	-	-
C32	14.35	9.98 / 9.57	-
T33	13.44	-	-
T34	13.05	-	-
C35	14.26	9.96 / 9.54	-
C36	-	-	-

4.2.2 Assignment of non-exchangeable resonances

Work towards the assignment of the non-exchangeable proton resonances began by examining the C/dU H5-H6 region of the TOCSY and NOESY spectra of the sample in D_2O solution. When the TOCSY was superimposed on to the TOCSY spectrum that was obtained for the hairpin (Figure 4.8), it could be seen that seven C/dU H5-H6 cross peaks overlaid well, with very little difference in the chemical shifts. These cross peaks were therefore assigned as the cytidine and deoxyuridine residues within the pyrimidine strand of the hairpin. An additional six TOCSY cross peaks were expected in this region for the triplex, corresponding to the six cytidine residues within the third strand of the triplex. Two cross peaks were clearly resolved, however the remainder were overlapped.
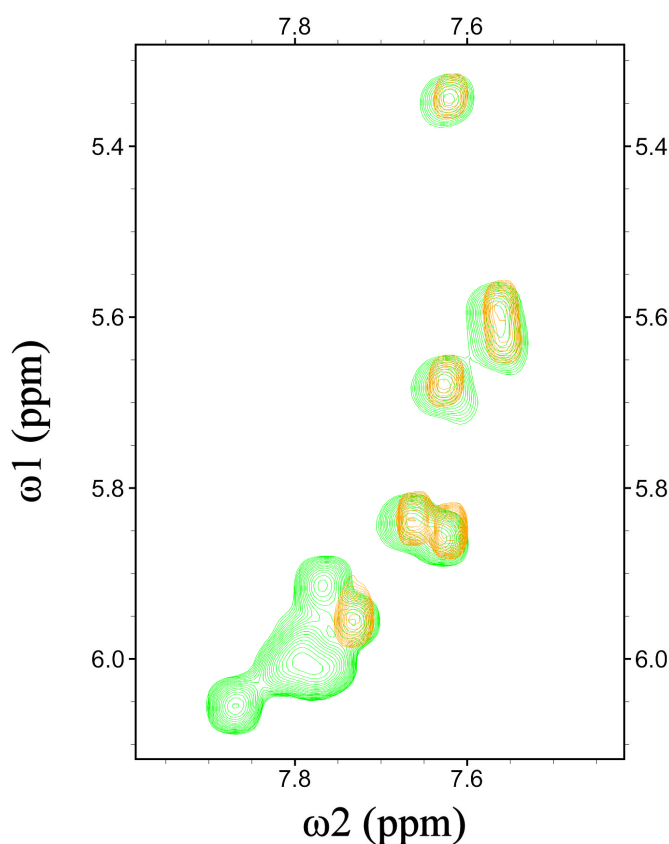
Figure 4.8. Expansion of the TOCSY spectrum for the hairpin (yellow) and non-modified triplex (green) showing the C/dU H5-H6 region.

The H5/H1′-H6/H8 region of the hairpin and triplex NOESY spectra also showed good convergence (Figure 4.9), with few chemical shift differences between the purine H1′-H8 and pyrimidine H1′-H6 cross peaks. This facilitated the assignment of all resonances corresponding to the Watson-Crick base paired strands of the triplex, and assignments were confirmed via imino-H8 assignments made on the (9:1) H_2O:D_2O NOESY spectrum for the triplex.
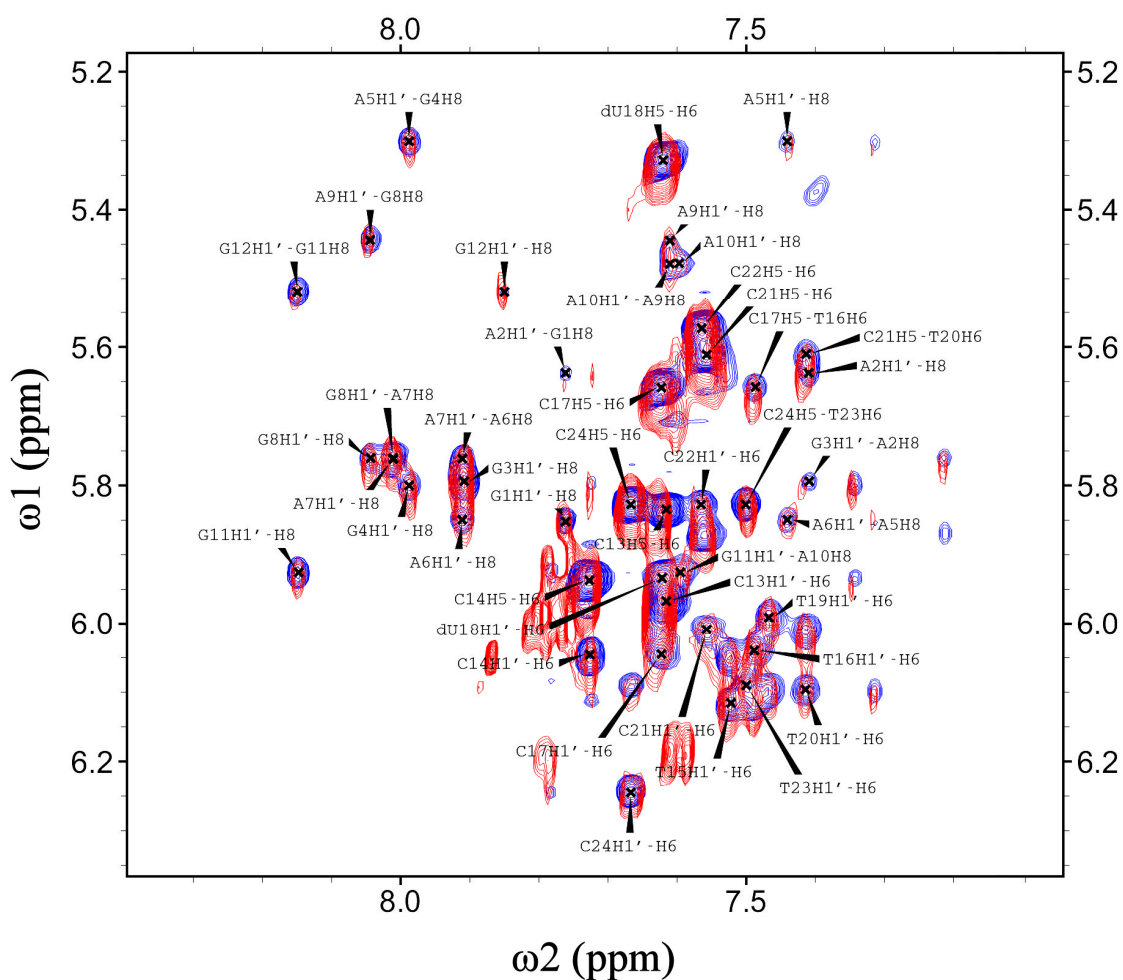
Figure 4.9. Expansions of the NOESY spectrum for the hairpin (blue) and non-modified triplex (red) showing the C/dU H5-H6 region.

A starting point for the third strand assignments was provided by examining both the C H5-H6 and thymidine Me-H6 regions of the TOCSY and NOESY spectra. Assignments for the Me-H6 cross peaks corresponding to the Watson-Crick base paired residues were again based on assignments for the hairpin, and confirmed via triplex Me-imino cross peak assignments. As with the C H5-H6 region there was considerable overlap between the third strand T Me-H6 cross peaks rendering assignments difficult. However, since the chemical shift of T34 H6 was known from the assignment of the T34 imino-H6 cross peak, the T34 Me-H6 cross peak could be identified. The C35 H5-H6 nOe cross peak could also be assigned based on the H5 chemical shift which had been identified via an intra-residue connection to its amino protons. This also facilitated the assignment of C36 H5-H6, which was attributed to the only remaining cytidine residue which did not have an nOe from its H6 proton to a T methyl.

Three inter-residue nOe connections which connected three thymidine methyl protons to the H6 protons on adjacent 5′ cytidine residues were observed. One of these T residues also exhibited an inter-residue connectivity between its H6 proton and an adjacent C H5, and so could be attributed to T26 - the only third strand T which has both 3′ and 5′ cytidine neighbours. C25 and C27 H5-H6 cross peaks could then be assigned via the sequential nOes to T26 (Figure 4.10). Due to the extensive overlap in the C H5-H6 region, no sequential connection between C27 H6 and C28 H5 was observed, therefore it was assumed that this nOe must be concealed underneath nearby overlapping cross peaks. Based on this assumption, the C28 H5-H6 interaction was assigned as was C32, which was attributed to the remaining cytidine intra-residue cross peak in this region.
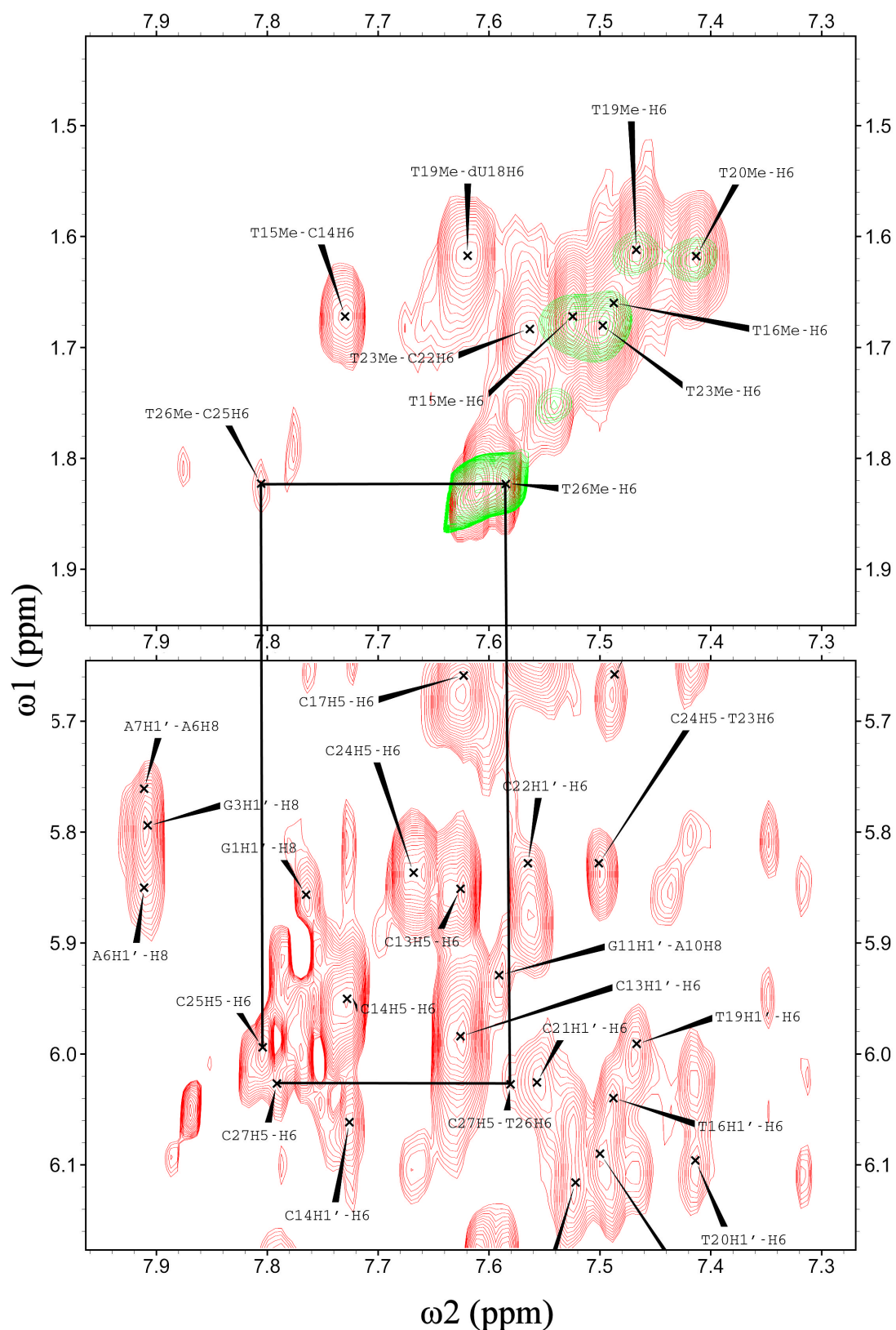
Figure 4.10. Expansions showing the Me-H6 region of the TOCSY (green) and NOESY (red) spectra, and the C/dU H5-H6 region of the NOESY spectrum for the non-modified triplex, highlighting sequential nOe connections between residues C25, T26 and C27.

The T29 Me-H6 cross peak could then be identified via a sequential connection to the H6 proton of C28, and T30 was assigned via a weak nOe to T29. The methyl chemical shift of T33 was obtained from the inter-residue connection to the H6 of C32; however the Me-H6 was not assigned due to extensive overlap in the region in which the H6 proton resonated. Figure 4.11 shows the extent to which Me-H6 assignments were made for the native triplex.
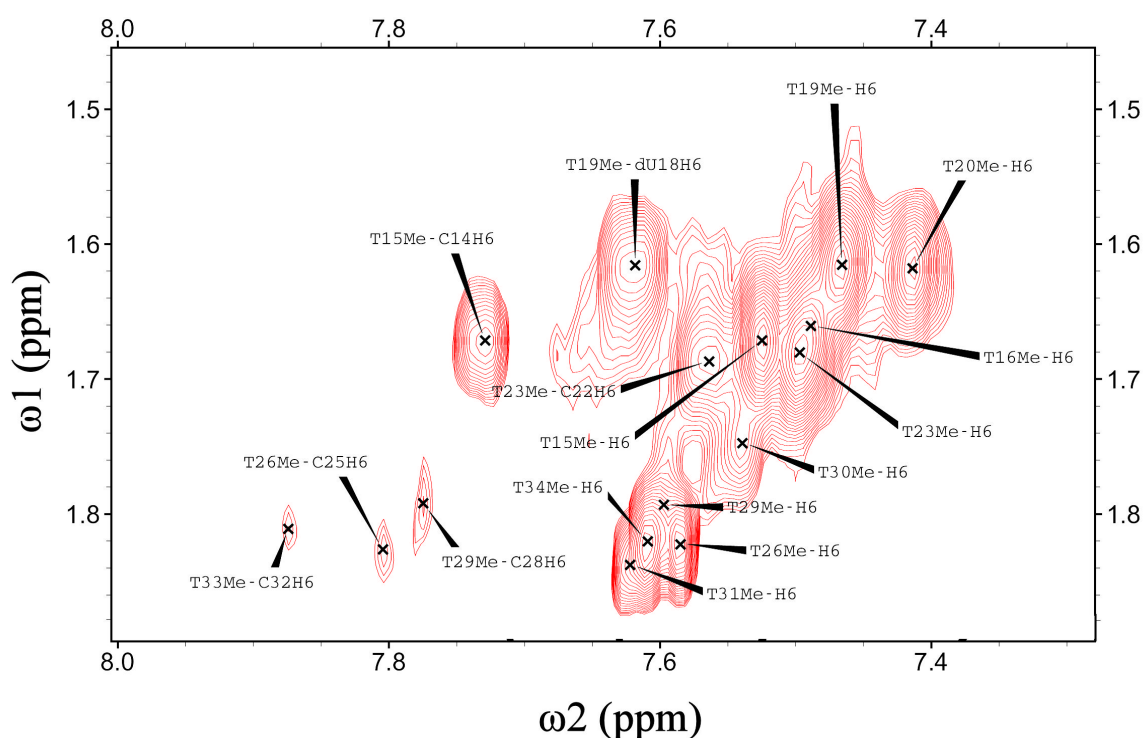
Figure 4.11. Fully assigned Me-H6 region of the NOESY spectrum for the non-modified triplex.

Once the base and H1′ proton assignment was completed as fully as possible, correlations to the remaining sugar protons could be identified. This facilitated the assignments of nOe cross peaks in the remaining regions of the NOESY spectrum, several of which are shown in Figures 4.12 and 4.13. The chemical shifts of all assigned protons for the native triplex are presented in Table 4.2.
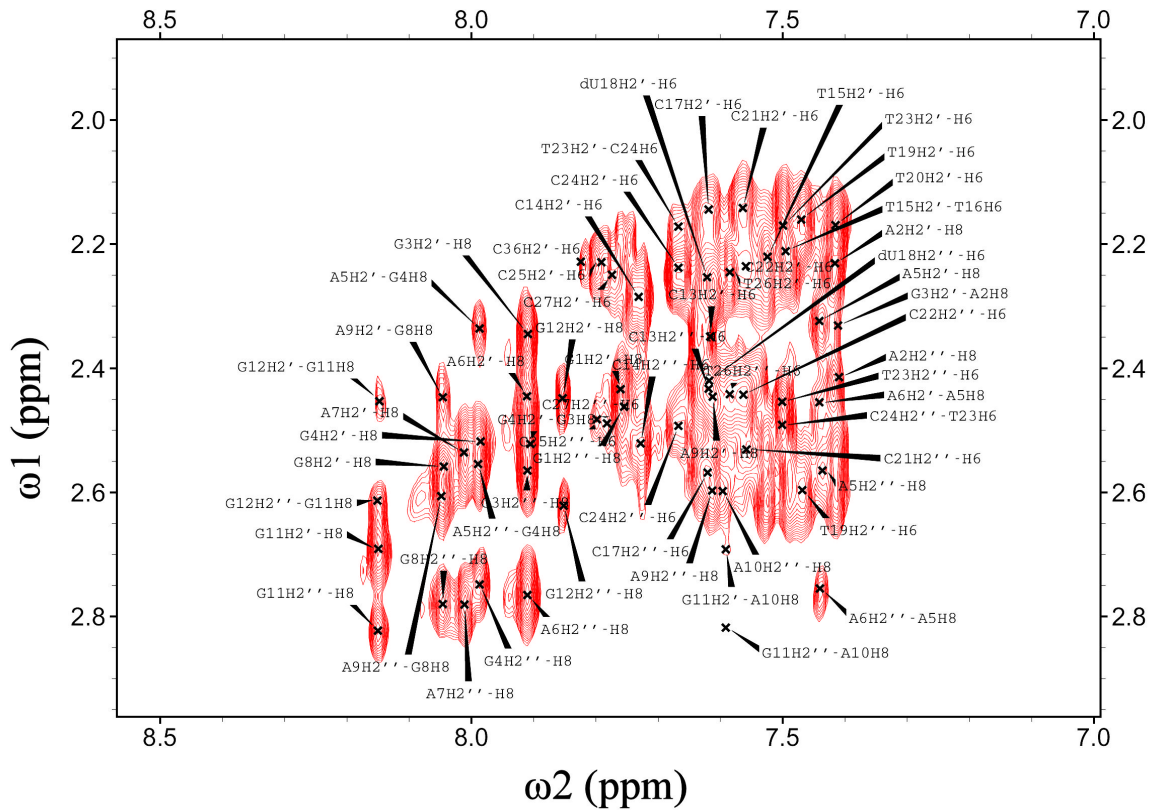
Figure 4.12. Expansion of the NOESY spectrum showing H2′/H2″-H6 correlations.
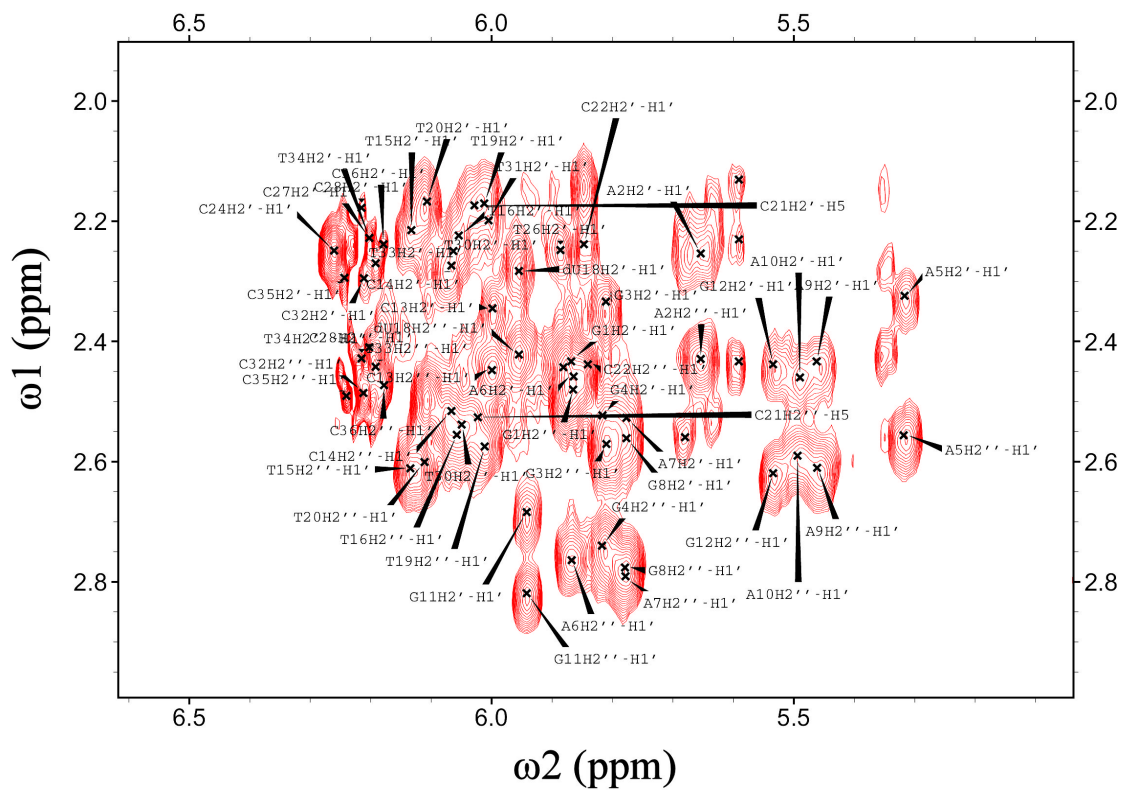
Figure 4.13. Expansion of the NOESY spectrum showing H2′/H2″-H1′ correlations.

Table 4.2. Chemical shifts (in ppm) of all non-exchangeable protons for the non-modified triplex at 20 °C, referenced to H₂O at 4.8 ppm.

Residue	H8/H6	Me/H5	H1′	H2′	H2″	H3′	H4′	H5′/H5″
G1	7.76	-	5.86	2.42	2.46	-	4.14	4.06, 4.03
A2	7.41	-	5.64	2.24	-	-	-	4.08
G3	7.91	-	5.80	-	2.57	4.95	4.31	4.11
G4	7.99	-	5.80	2.53	2.74	4.98	-	4.14, 4.08
A5	7.44	-	5.30	2.33	2.55	-	-	-
A6	7.91	-	5.85	2.45	2.76	4.97	4.34	-
A7	8.01	-	5.78	-	-	5.00	4.34	-
G8	8.05	-	5.72	2.57	2.78	5.00	4.36	4.12, 4.10
A9	7.61	-	5.44	2.44	2.61	-	-	-
A10	7.60	-	5.48	2.46	2.59	-	-	-
G11	8.15	-	5.94	2.69	2.82	4.99	-	4.12
G12	7.85	-	5.52	2.44	2.62	-	-	3.64
C13	7.63	5.86	5.98	2.35	2.43	-	4.04	4.01, 3.85
C14	7.73	5.95	6.06	2.28	2.52	4.74	4.22	4.05, 3.99
T15	7.52	1.67	6.12	2.21	2.61	4.86	4.21	-
T16	7.49	1.67	6.04	2.22	2.57	-	-	-
C17	7.63	5.67	6.03	2.14	2.56	4.81	4.09	-
dU18	7.62	5.34	5.93	2.26	2.42	4.85	-	-
T19	7.47	1.61	5.99	2.16	2.57	-	4.21	4.15, 4.10
T20	7.41	1.62	6.11	2.16	2.61	-	4.20	4.16
C21	7.56	5.61	6.02	2.13	2.13	-	4.18	4.10
C22	7.56	5.59	5.87	2.24	2.44	-	-	-
T23	7.50	1.68	6.09	2.17	2.45	-	4.12	4.06
C24	7.67	5.84	6.25	2.24	2.24	4.54	4.13	3.40, 3.99
C25	7.81	6.00	6.20	2.23	2.48	-	-	3.77
T26	7.59	1.82	6.19	2.25	2.44	-	4.08	-
C27	7.79	6.03	6.12	2.24	2.49	-	4.34	-
C28	7.77	6.01	-	2.24	-	-	-	-
T29	7.60	1.80	-	2.21	2.38	4.54	-	-
T30	7.54	1.74	6.05	2.25	-	-	4.14	-
T31	7.62	1.84	6.01	2.22	-	-	4.22	3.63
C32	7.87	6.05	6.21	2.30	2.49	-	4.26	-
T33	-	1.81	6.19	2.27	2.44	4.52	4.27	-
T34	7.61	1.82	6.19	2.23	-	-	-	-
C35	7.79	5.91	6.24	2.29	2.49	4.80	-	-
C36	7.82	6.01	6.20	2.24	2.47	4.55	4.27	-

4.3 NMR assignments for the modified triplex

NMR assignments for the 3′-*S*-phosphorothiolate modified triplex followed the same procedure as for the non-modified equivalent. The sequence and numbering for the modified system is shown in Scheme 4.2.

$$5'-C_{25}^+\ T_{26}\ C_{27}^+\ C_{28}^+\ T_{29}\ \boxed{T_{S\,30}\ T_{31}}\ C_{32}^+\ \boxed{T_{S33}T_{34}}\ C_{35}^+\ C_{36}^+-3'$$

$$5'-G_1\ A_2\ G_3\ G_4\ A_5\ A_6\ A_7\ G_8\ A_9\ A_{10}\ G_{11}\ G_{12}$$

$$3'-C_{24}\ T_{23}\ C_{22}\ C_{21}\ T_{20}\ T_{19}\ dU_{18}\ C_{17}\ T_{16}\ T_{15}\ C_{14}\ C_{13}$$
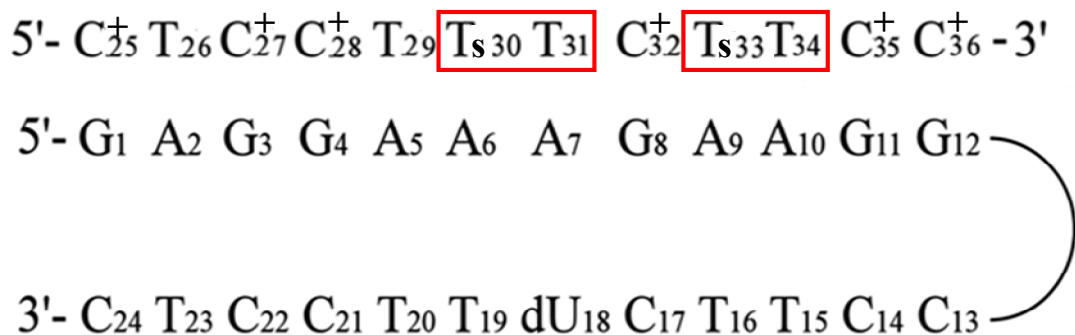
Scheme 4.2. The sequence and numbering of the modified triplex under investigation. 3′-*S*-phosphorothiolate linkages are highlighted red.

Triplex formation was confirmed by the presence of two sets of broad downfield shifted resonances in the one dimensional spectrum which are characteristic of imino and amino protons from the protonated third strand cytidine residues (Figure 4.14).
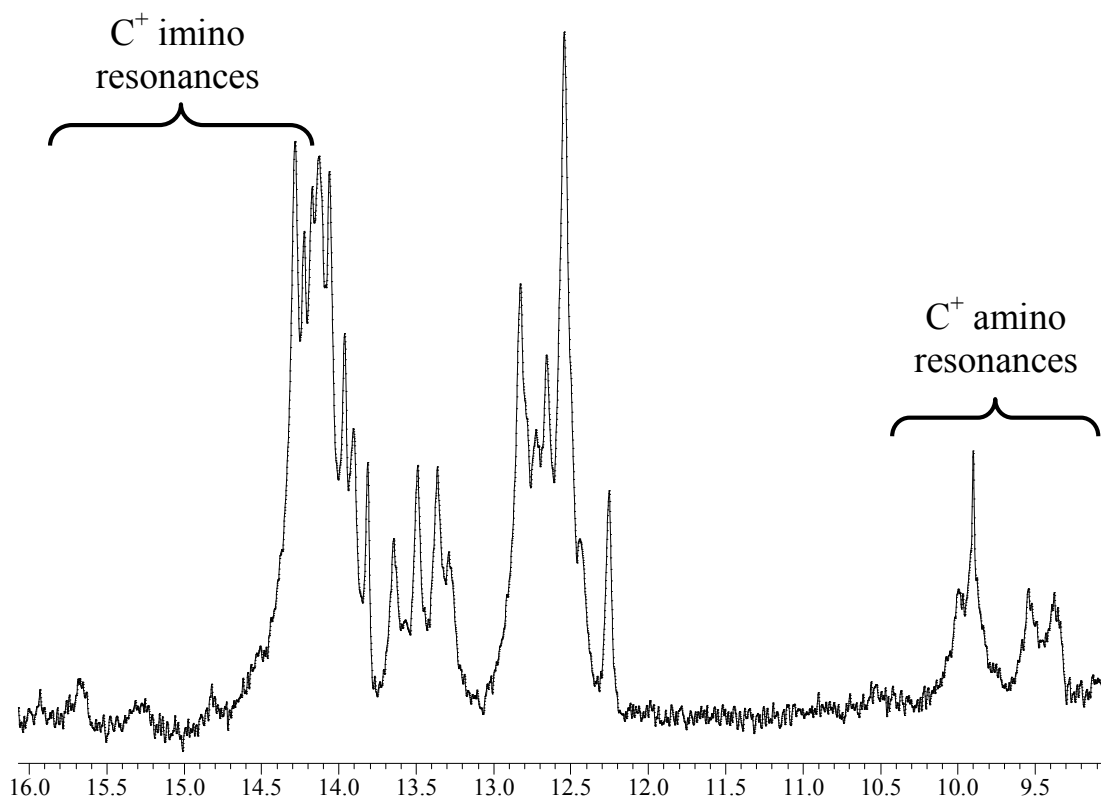
Figure 4.14. Expansion of the one dimensional ^{1}H spectrum of the modified triplex at pH 5, recorded in a 9:1 H$_2$O:D$_2$O solution mix at 283 K. The lowfield imino and C$^+$ amino resonances that are characteristic of a parallel triplex are highlighted.

4.3.1 Assignment of exchangeable resonances

The first step in assigning the exchangeable protons for the 3′-*S*-phosphorothiolate modified triplex was to identify nOe connections in the imino and amino regions of the NOESY spectrum recorded in a 9:1 H$_2$O:D$_2$O solvent mix. As with the non-modified triplex, the same pattern of sequential connections between the imino protons on the Watson-Crick base paired strands could be identified, facilitating the assignment of these protons (Figure 4.15).
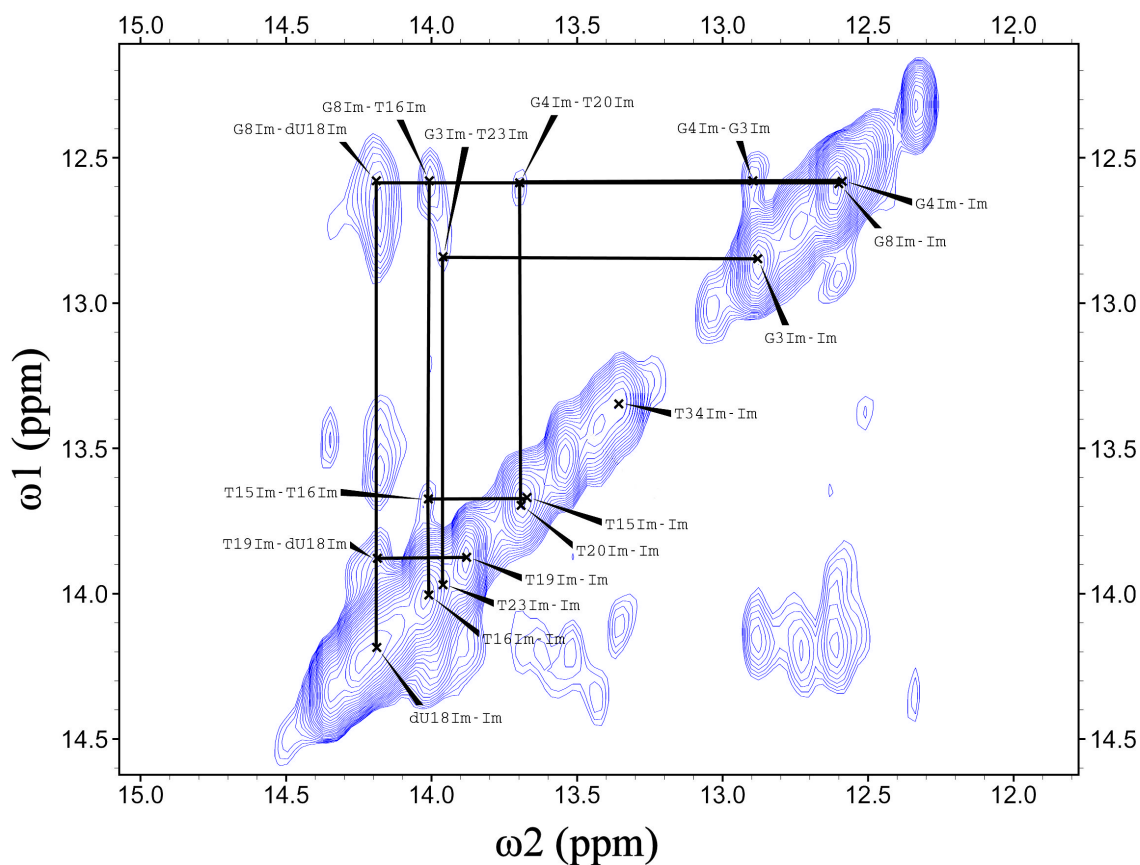
Figure 4.15. Expansion of the NOESY spectrum for the modified triplex in a (9:1) H₂O:D₂O solution, showing sequential connections between the Watson-Crick bonded imino protons.

Imino protons corresponding to the third strand cytidine residues were identified via connections to their amino protons, and three of these could be unambiguously assigned as C28, C32 and C35 as they exhibited sequential assignments to thymidine imino protons, which are highlighted in Figure 4.16. Based on these nOe connections, T29, T31, Ts33 and T34 imino protons were also assigned. T26 and Ts30 imino protons were assigned based on nOe connections to their methyl protons and assignments were confirmed by comparing to assignments made for the non-modified triplex NOESY spectrum.
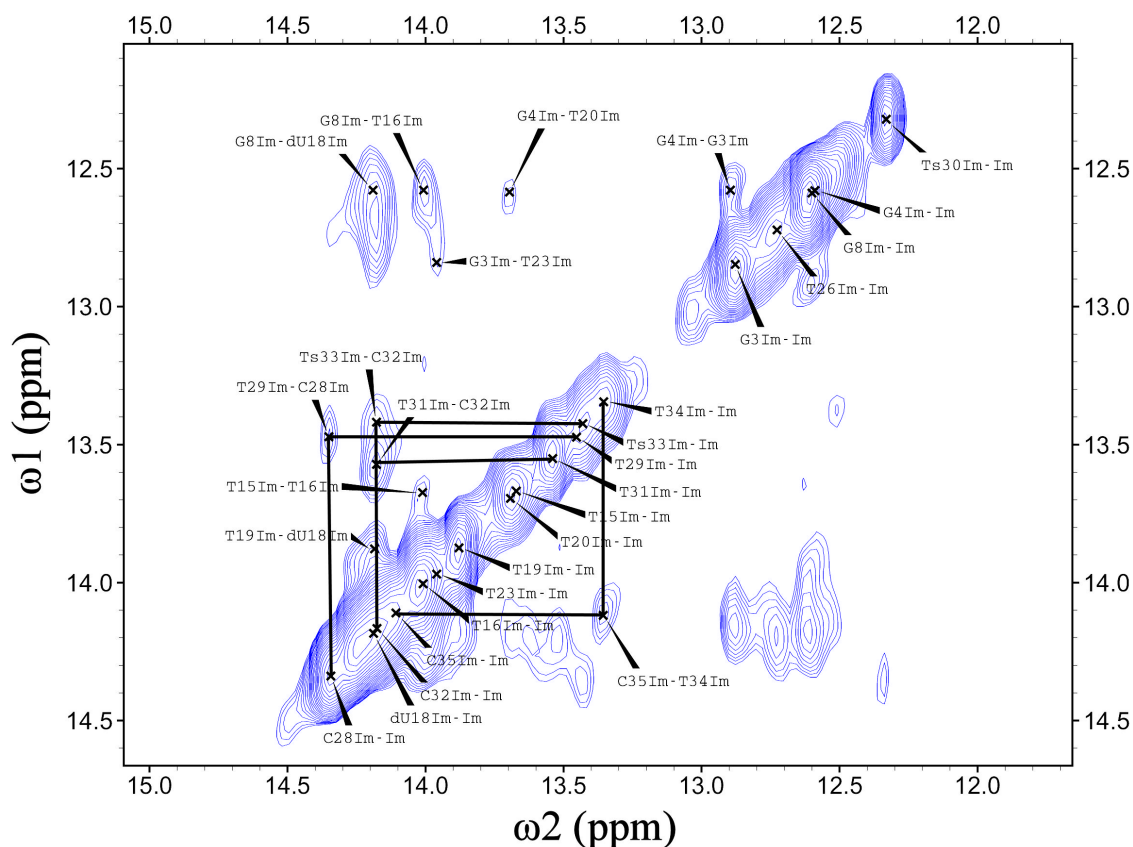
Figure 4.16. Expansion of the NOESY spectrum for the modified triplex in a (9:1) H₂O:D₂O solution, highlighting some connections between imino protons within the Hoogsteen bonded third strand.

Once the imino-imino proton region of the NOESY spectrum was assigned as fully as possible, assignments could be transferred to the amino/aromatic-imino proton region, facilitating the assignments of the amino protons from the Watson-Crick base paired strands of the triplex and completing the assignment of the exchangeable resonances. Connections between the imino protons to several H2, H6 and H8 resonances were also observed. Cross peaks in this region were extremely broad and due to extensive overlap several resonances could not be identified; the extent to which assignments were completed is highlighted in Figure 4.17. As with the non-modified triplex, this region can provide a useful starting point for the assignment of the non-exchangeable proton resonances.
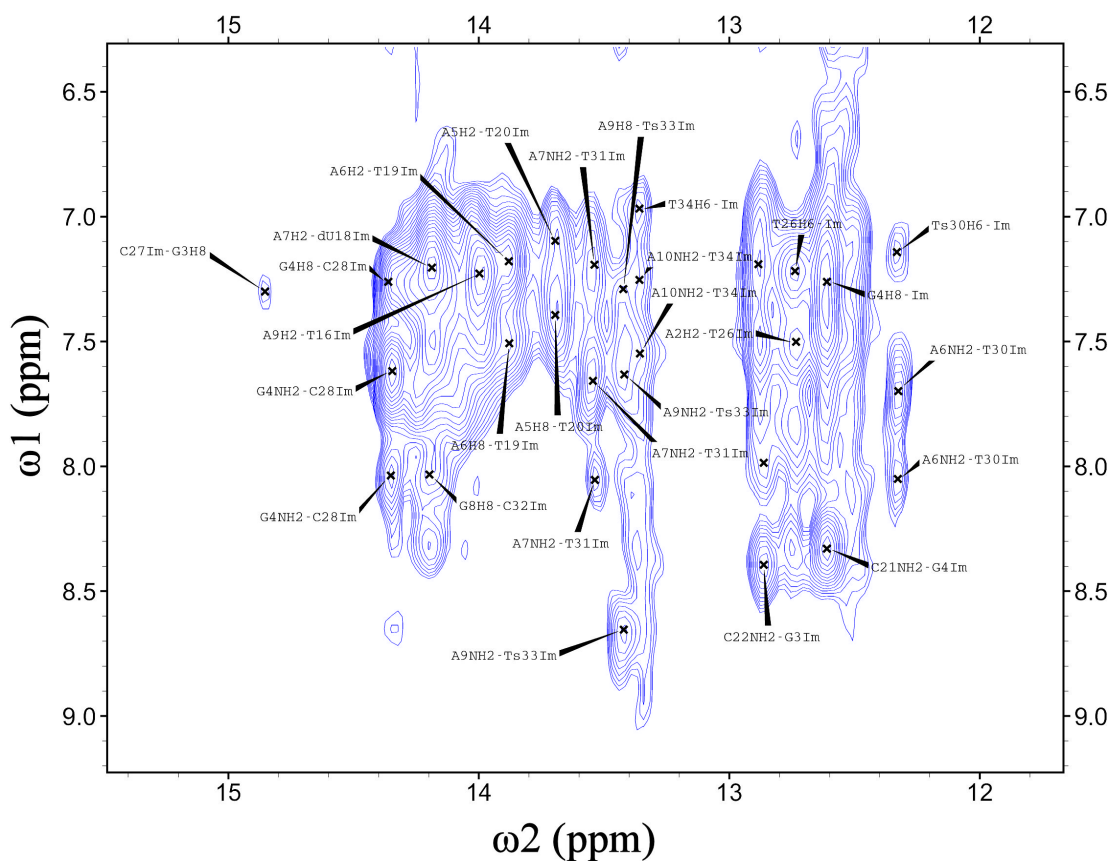
Figure 4.17. Expansion of the NOESY spectrum of the modified triplex in a (9:1) H$_2$O:D$_2$O solution, showing nOe connections between imino and amino/aromatic protons.

Connections between the pyrimidine imino protons from the Hoogsteen strand to the H2′ and H2″ protons on 5′-neighbouring purine residues were again observed and could be assigned (Figure 4.18), providing further evidence for the formation of the modified triplex and another potential starting point for non-exchangeable assignments.
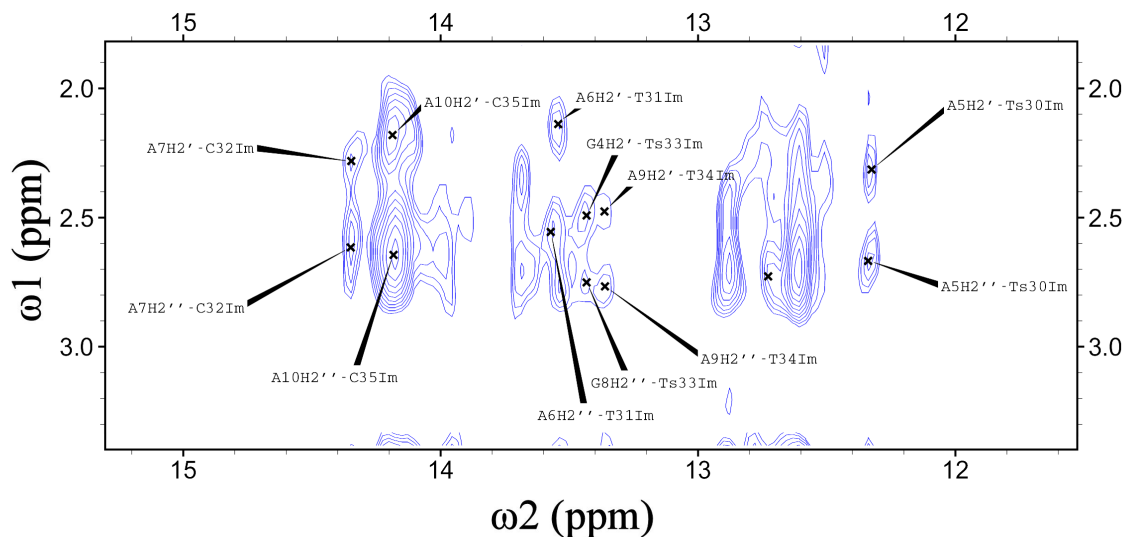
Figure 4.18. Expansion of the NOESY spectrum of the modified triplex in a (9:1) H₂O:D₂O solution, showing connections between third strand imino protons and 5′-neighbouring purine H2′/H2″ protons.

Chemical shifts of all assigned labile and adenosine H2 protons for the 3′-*S*-phosphorothiolate modified triplex are given in Table 4.3. The majority of imino protons were assigned with the exception of those corresponding to the residues at either end of the oligonucleotide.

Table 4.3. Chemical shifts (in ppm) of all exchangeable and H2 protons for the modified triplex at 2 °C, referenced to H₂O at 4.95 ppm.

Residue	Imino	Amino	H2
G1	-	-	-
A2	-	-	-
G3	12.88	-	-
G4	12.59	8.04 / 7.62	-
A5	-	-	7.10
A6	-	8.05 / 7.70	7.18
A7	-	8.05 / 7.66	7.21
G8	12.60	-	-
A9	-	8.56 / 7.63	7.24
A10	-	7.57 / 7.26	-
G11	-	-	-
G12	-	-	-
C13	-	-	-
C14	-	-	-
T15	13.67	-	-
T16	14.01	-	-
C17	-	-	-
dU18	14.19	-	-
T19	13.88	-	-
T20	13.69	-	-
C21	-	8.33	-
C22	-	8.39	-
T23	13.95	-	-
C24	-	-	-
C25	-	-	-
T26	12.73	-	-
C27	14.85	9.98	-
C28	14.34	-	-
T29	13.45	-	-
Ts30	**12.33**	**-**	**-**
T31	13.54	-	-
C32	14.18	9.96 / 9.57	-
Ts33	**13.43**	-	-
T34	13.35	-	-
C35	14.11	-	-
C36	-	-	-

4.3.2 Assignment of non-exchangeable resonances

The thymidine Me-H6 region of the NOESY spectrum provided a useful starting point for the modified triplex non-exchangeable proton assignments. There was much less overlap in this region than for the unmodified version, and resonances were better resolved. However, no cross peaks were observed in this region in the TOCSY spectrum, as thymidine intra-base TOCSY connections are typically weak, and the modified triplex sample was dilute compared to the non-modified triplex sample, resulting in a poorer signal:noise ratio.

The two modified residues (Ts30 and Ts33) were identified by their H3′-H6 nOe connections, as the H3′ protons were shifted approximately 1 ppm upfield from their expected chemical shifts. These are shown in Figure 4.19. From here the corresponding intra-residue Me-H6 cross peaks could be identified and assignments were made based on sequential nOes in this region, which are highlighted in Figure 4.20. One modified residue was connected to two neighbouring thymidines in a 'TTT' sequence; these were assigned as T29, Ts30 and T31, and the other was connected to one neighbouring thymidine in the 3′ direction and which were assigned as Ts33 and T34.
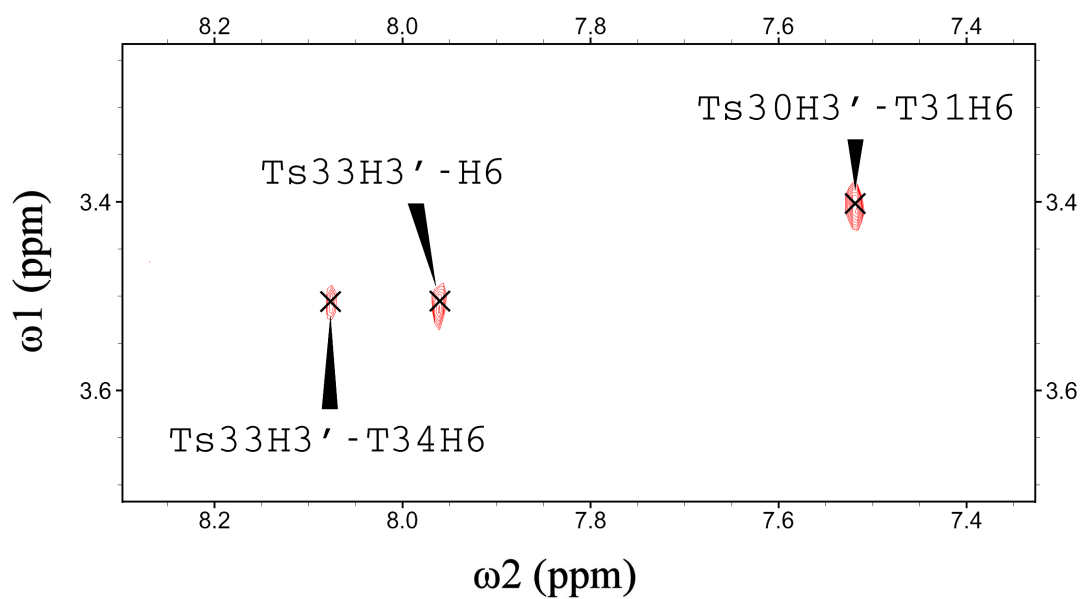
Figure 4.19. Expansion of the NOESY spectrum for the modified triplex, showing connections between the upfield shifted Ts30 and Ts33 H3′ and H6 proton resonances.
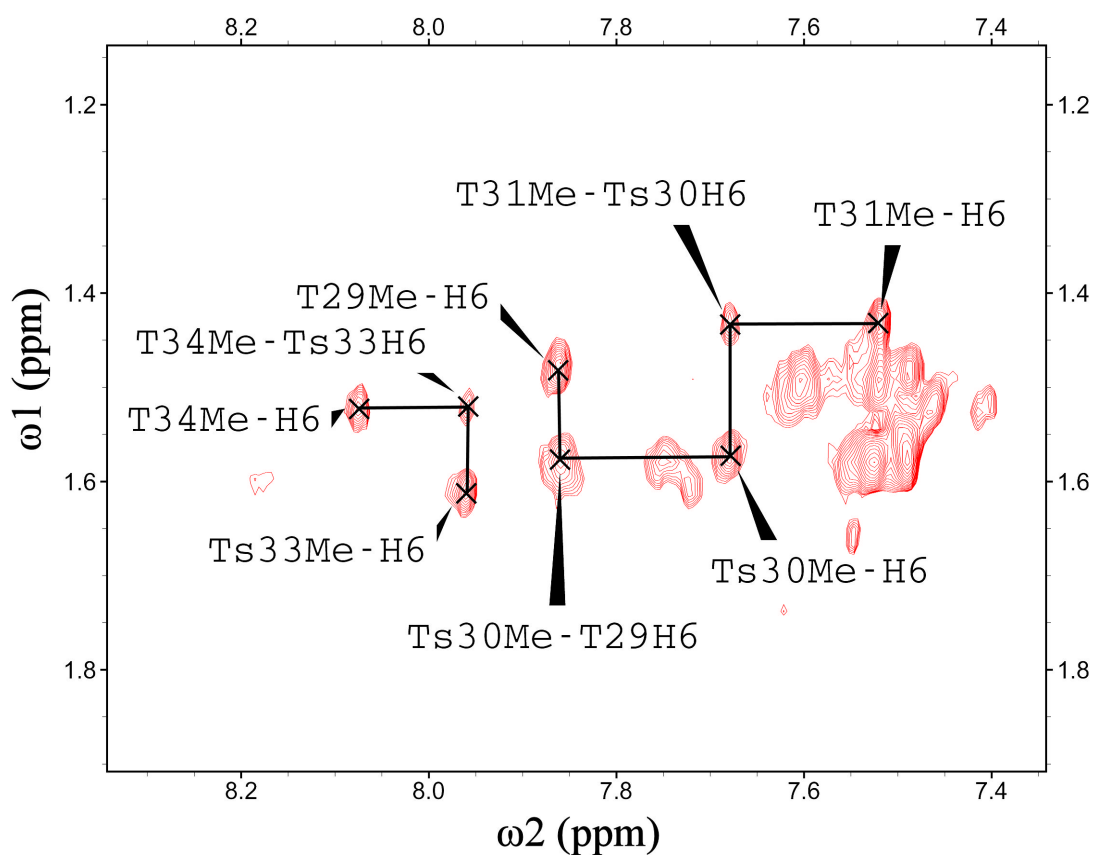
Figure 4.20. Expansion of the Me-H6 region of the NOESY spectrum for the modified triplex, highlighting sequential nOe connections between neighbouring thymidine residues in the third strand.

From the assignment of sequential thymidines and the identification of intra-base cross peaks from the TOCSY spectrum, the remaining pyrimidine Me/H5-H6 cross peaks were assigned following the same methodology that was used for the assignment of the native triplex. Several of the purine H8 protons could be identified from the imino-H8 assignments; from here H1′-H8 assignments could be made by following the sequential chain of nOes.

Once the sequential base and H1′ protons had been unambiguously assigned, the remaining sugar protons were assigned using the TOCSY and NOESY spectra. The modified triplex was assigned as fully as possible, however some protons could not be confidently assigned due to significant overlap between several cross peaks, which is usual for an oligonucleotide of this size. Chemical shifts for the assigned protons are presented in Table 4.4.

Table 4.4. Chemical shifts (in ppm) of all non-exchangeable protons for the modified triplex at 20 °C, referenced to H₂O at 4.8 ppm.

Residue	H8/H6	Me/H5	H1′	H2′	H2″	H3′	H4′	H5′/H5″
G1	7.89	-	5.98	2.35	2.49	-	-	-
A2	7.49	-	5.72	2.10	2.38	4.80	-	-
G3	7.91	-	5.72	2.26	2.52	4.87	4.51	-
G4	8.00	-	5.82	2.45	2.70	4.83	4.26	-
A5	7.45	-	-	2.39	-	-	-	-
A6	7.92	-	5.90	2.40	2.67	4.82	4.27	-
A7	8.02	-	5.74	2.45	2.71	4.86	-	-
G8	8.04	-	5.78	2.50	2.70	-	-	4.07
A9	7.63	-	5.59	-	-	-	-	-
A10	7.45	-	5.58	-	-	4.42	-	-
G11	8.14	-	5.92	2.59	2.73	-	-	-
G12	7.87	-	5.66	2.39	2.53	-	4.21	3.99
C13	7.75	5.87	5.98	2.30	-	-	4.22	4.12
C14	7.75	5.84	6.15	2.19	2.42	-	4.22	4.03, 3.90
T15	7.53	1.57	6.02	-	-	-	-	-
T16	7.49	1.48	5.94	-	2.44	-	4.26	-
C17	7.57	5.53	5.90	2.26	2.45	4.71	-	3.89
dU18	7.63	5.25	5.93	2.19	2.47	-	4.13	4.04
T19	7.47	1.52	5.90	2.08	2.51	-	-	-
T20	7.49	1.57	6.02	-	2.65	4.80	-	4.08
C21	7.54	5.44	5.74	-	2.27	-	-	-
C22	7.61	5.40	5.81	-	2.54	4.57	-	4.09
T23	7.53	1.50	5.90	2.11	2.44	-	-	-
C24	7.98	5.77	6.15	2.41	2.30	4.57	4.24	4.13, 4.02
C25	7.86	6.10	6.10	2.14	2.34	4.41	4.26	4.06, 3.88
T26	7.86	1.58	5.95	2.37	2.45	-	-	-
C27	7.67	5.79	5.91	2.13	2.53	4.80	4.23	4.08, 4.02
C28	7.66	5.74	5.89	2.19	2.31	-	-	-
T29	7.86	1.48	5.77	2.31	2.52	4.40	4.22	4.08
Ts30	**7.68**	**1.57**	**6.03**	**2.22**	**2.34**	**3.41**	**4.02**	-
T31	7.52	1.44	6.03	2.05	2.54	-	4.28	4.09, 4.04
C32	7.74	5.62	6.11	2.19	2.43	-	4.45	4.22
Ts33	**7.97**	**1.61**	**-**	**2.40**	**2.52**	**3.51**	**4.24**	**4.07**
T34	7.83	1.52	6.20	2.40	2.52	-	-	4.05, 4.03
C35	7.86	5.84	6.14	2.38	2.60	4.38	4.25	4.05
C36	7.79	5.74	-	-	2.51	4.62	4.25	4.09

4.4 Qualititative structural analysis

There was significantly more crowding and overlapping of nOe cross peaks in the NOESY spectra for the triplexes, which both comprise 50 % more residues compared to the hairpin. This limited the qualitative analysis which could be performed. The glycosidic bond orientation was not determined due to the lack of H3′-H6 cross peaks that were apparent on both NOESY spectra and the poor resolution of H1′-H6 cross peaks, particularly for the non-modified triplex NOESY spectrum.

4.4.1 Chemical shift analysis

The chemical shifts (δ) of protons are dominated by the electron densities around the nuclei and the electronegativity of neighbouring atoms/groups. Any change in δ between specific protons in the non-modified and modified triplexes may indicate a change in the local electronic environment of that proton. Significant changes (of 0.05 ppm or greater) for assigned protons in both the unmodified and modified triplex NMR spectra are highlighted in Table 4.5, and illustrated graphically for the base H6/H8, H1′ and H3′ protons in Figure 4.21.

Table 4.5. Chemical shift differences (Δδ) between protons within the native and modified triplexes, given as Δδ = δ$_{(modified)}$ − δ$_{(native)}$.

Residue	Observed shift changes (ppm)
G1	**H8** 0.12, **H1′** 0.10, **H2′** − 0.07
A2	**H8** 0.08, **H1′** 0.08, **H2′** − 0.13
G3	**H1'** − 0.08, **H3'** − 0.08, **H4'** 0.19
G4	**H1′** 0.02, **H2′** − 0.07, **H3'** − 0.15
A5	**H2′** 0.06,
A6	**H2** 0.17, **H1′** 0.05, **H2′** − 0.06, **H2″** − 0.09, **H3'** − 0.15, **H4'** − 0.07
A7	**H2** − 0.06, **H3'** − 0.14
G8	**H1′** 0.06, **H2′** − 0.06, **H2″** − 0.08
A9	**H1′** 0.14
A10	**H8** − 0.14
G11	**H2′** − 0.10, **H2″** − 0.09
G12	**H1′** 0.14, **H2″** − 0.09
C13	**H6** 0.11, **H4'** 0.18
C14	**H5** − 0.11, **H1′** 0.09, **H2′** − 0.09, **H2″** − 0.10
T15	**Me** − 0.10, **H1′** − 0.10, **Im** − 0.13
T16	**Me** − 0.18, **H1′** − 0.10, **H2″** − 0.13, **Im** − 0.1
C17	**H6** − 0.06, **H5** − 0.14, **H1′** − 0.13, **H2′** 0.12, **H2″** − 0.11, **H3'** − 0.09
dU18	**H5** − 0.09, **H2′** − 0.07, **Im** 0.08
T19	**Me** -0.09, **H1′** -0.09, **H2′** -0.07, **H2″** -0.06, **Im** 0.13
T20	**H6** 0.08, **H1′** -0.09, **Im** − 0.2
C21	**H5** − 0.17, **H1′** − 0.28, **H2″** 0.14
C22	**H6** 0.05, **H5** − 0.19, **H1′** − 0.06, **H2″** 0.11
T23	**Me** − 0.18, **H1′** − 0.19, **H2′** − 0.06, **Im** − 0.14
C24	**H5** − 0.06, **H1′** − 0.09, **H2′** − 0.17, **H2″** 0.07, **H4'** 0.12
C25	**H6** 0.06, **H5** 0.11, **H1′** − 0.10, **H2′** − 0.09, **H2″** − 0.14
T26	**Me** − 0.24, **H1′** − 0.24, **H2′** 0.12
C27	**H6** − 0.13, **H5** − 0.24, **H1′** − 0.21, **H2′** − 0.11, **H4'** − 0.11, **Im** 0.14
C28	**H6** − 0.11, **H5** − 0.27, **Im** − 0.46
T29	**Me** − 0.32, **H2′** 0.11, **H2″** 0.15, **H3'** − 0.14, **Im** 0.11
T30/Ts30	**H6** 0.14, **H5** − 0.17, **H1′** − 0.02, **H4'** − 0.12, **Im** − 0.11
T31	**H6** − 0.10, **H2′** − 0.17, **H4'** 0.07, **Im** − 0.08
C32	**H6** − 0.13, **H5** − 0.43, **H1′** − 0.11, **H2′** − 0.11, **H2″** − 0.06, **H4′** 0.19, **Im** − 0.17
T33/Ts33	**Me** − 0.20, **H2′** 0.13, **H2″** 0.08, **H3'** − 1.01
T34	**H6** 0.22, **Me** − 0.29, **H2′** 0.17, **Im** 0.30
C35	**H6** 0.06, **H5** − 0.07, **H2′** 0.09, **H2″** 0.10, **Im** − 0.15
C36	**H5** − 0.27, **H3'** 0.07
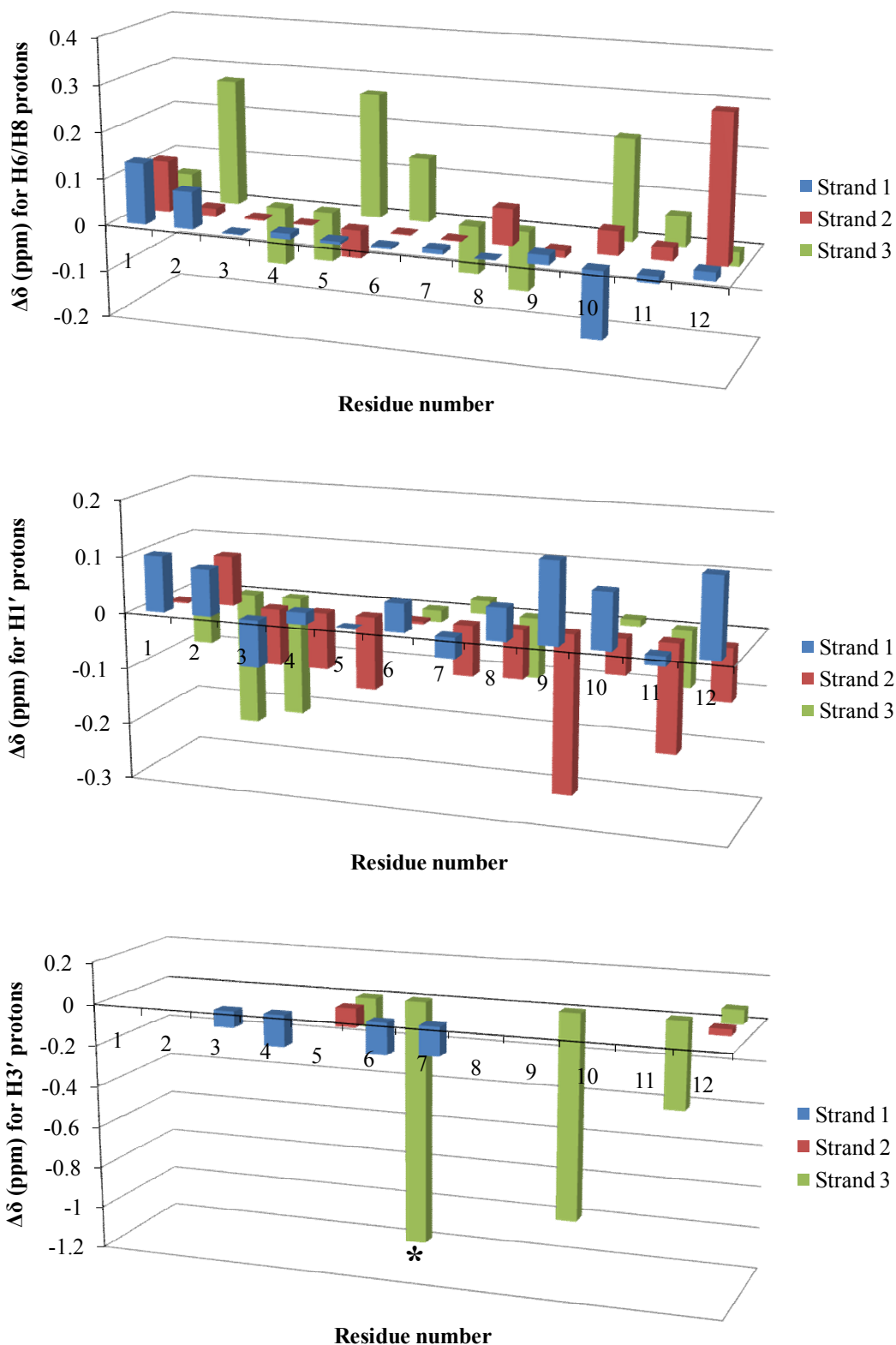
Figure 4.21. Chemical shift differences (Δδ) between the modified and non-modified triplexes base H6/H8, H1′ and H3′ protons, given as $\Delta\delta = \delta_{(modified)} - \delta_{(native)}$.

 * T30/Ts30 H3′ Δδ was estimated as δ could not be determined for the T30 H3′ proton in the non-modified triplex NOESY spectrum.

The largest observed chemical shift difference (Δδ) corresponds to the H3′ proton between T33 and the modified residue Ts33. No Δδ can be reported for the H3′ proton between T30 and the other modified nucleotide, Ts30, as this proton could not be unambiguously assigned in the non-modified triplex spectra. However, the most upfield H3′ proton that was apparent on the native triplex NOESY spectrum resonated at 4.522 ppm, therefore since the H3′ of Ts30 in the modified triplex resonated at 3.407 it can be assumed that this proton experienced a significant change in chemical shift of at least – 1.12 ppm.

This is to be expected as the H3′ proton on these nucleotides is only two bonds away from the modification, and so the δ difference is a direct electronic effect of the substitution of oxygen for sulfur. The decreased electronegativity of the sulfur atom compared to oxygen reduces the deshielding of the H3′ proton causing it to resonate at a lower chemical shift. This effect may also explain the difference observed between the H4′ proton of T30 and Ts30, however electronic effects are short range and so may not explain the shift differences associated with the remaining protons on the modified nucleotides or protons on other residues.

Chemical shift differences in the base protons may be explained by changes in ring current shifts due to base stacking changes[115]. The H6 protons of the modified residues were deshielded and the H2 proton of A6 which is adjacent to a modified residue was shielded; this is consistent with an increase in intra-strand base stacking between the modified and 3′ adjacent nucleotides[115, 154, 155].

4.4.2 Sugar pucker determination

The majority of inter-residue H2′-H6 cross peaks were not well resolved and so accurate cross peak heights could not be determined. Sugar puckers for the individual residues within the triplexes were therefore determined using the H1′-H2′/H2″ cross peak height measurements (as described in Section 2.4.1). Figure 4.22 shows a one dimensional slice through the NOESY spectrum for the modified triplex, highlighting the relative cross peak heights for one of the modified thymidine residues, Ts30. The H1'-H2" peak height is larger than the H1'-H2' peak height, indicating that the H1'-H2" distance is shorter and so the Ts30 sugar is predominantly north. The individual sugar puckers for every nucleotide within the non-modified and modified triplexes for which peak heights could be obtained are presented in Table 4.6.
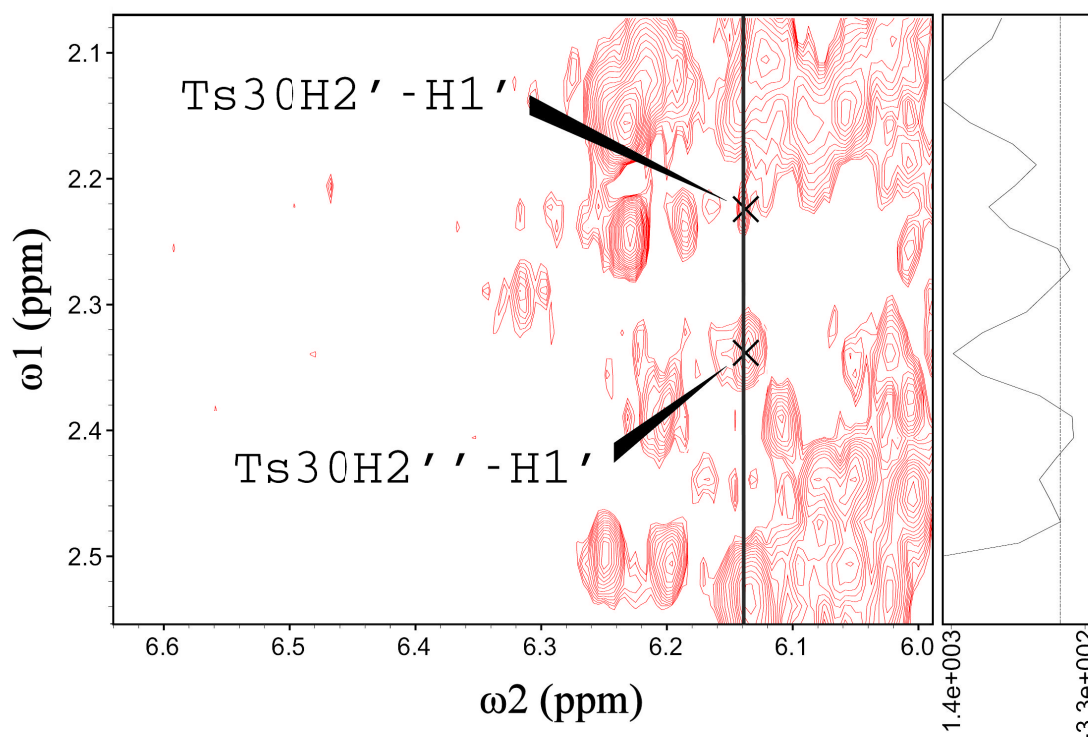
Figure 4.22. One dimensional slice through the modified triplex NOESY spectrum showing relative heights of the Ts30 H1'-H2' and H1'-H2'' cross peaks.

Table 4.6. Sugar puckers of the non-modified triplex residues deduced from the relative peak heights of H1′-H2′/H2″ NOESY cross peaks.

Residue	Non-modified triplex Sugar puckers	Modified triplex Sugar Puckers
G1	South	South
A2	South	South
G3	South	South
G4	South	-
A5	South	-
A6	South	-
A7	South	South
G8	South	South
A9	South	South
A10	South	-
G11	South	South
G12	South	-
C13	South	-
C14	South	-
T15	South	-
T16	-	South
C17	-	South
dU18	South	-
T19	South	South
T20	-	South
C21	-	South
C22	South	-
T23	-	South
C24	-	South
C25	-	South
T26	South	-
C27	South	South
C28	South	South
T29	-	South
T30/Ts30	-	**North**
T31	-	**North**
C32	South	South
T33/Ts33	South	**North**
T34	South	**North**
C35	South	-
C36	South	South

Although several sugar puckers could not be determined due to extensive overlap between cross peaks in the H2′/H2″-H1′ region of the NOESY spectrum, all of the evaluated sugars within the Watson-Crick hairpin portion of both triplexes tended to adopt a more south conformation. Likewise, only south sugar puckers were observed within the unmodified third strand. Residues in the modified third strand however, were found to contain a mixture of north and south sugar puckers, with the sugars attached to, and 3′ to, the modifications all preferentially adopting a north conformation. Unfortunately, one of the nucleotides 3′ to the modification (T31) could not be analysed with this method, as integration of the H2′-H1′ cross peak was impeded due to severe overlap with several other nOes. However, the view along the 1D slice implied that the T31 sugar prefers a north conformation, as the H2′-H1′ cross peak height appeared to be shorter than that of the T31 H2″-H1′.

4.5 Discussion

For both the native and modified versions, the exchangeable proton assignments provided evidence for triplex formation, indicating that under the experimental conditions (pH 5, 2 °C) the TFO binds to the hairpin by Hoogsteen base pairing.

Both the unmodified and modified versions were assigned as fully as possible using the procedure for a right handed, B-type helix. Severe overlap in several regions of the NOESY spectrum due to the size and homogeneity of the oligonucleotide prevented the assignments of several H3′/H4′/H5′/H5″ protons. Binding of the Hoogsteen strand appears to have little effect on the overall structure of the hairpin, as the base protons within the hairpin portion of the non-modified triplex were found to resonate at very similar chemical shifts to the protons of the hairpin itself. This

suggests that the third strand slots into the major groove, causing very little disruption to the hairpin helix. This is in agreement with a study by Sekharudu *et al*. in which a poly d(T) TFO was readily accommodated in the major groove of a poly d(A)-poly d(T) duplex with no perturbation[156].

Assignment of the labile protons was achieved for the majority of residues within the triplexes except for the terminal residues. For each triplex, two sets of nOe connections between hydrogen bonded imino and amino protons provided evidence for Watson-Crick base pairing between the two strands of the hairpin and Hoogsteen base pairing between the purine strand and the TFO.

Qualitative structural analysis was limited due to the extent of overlap in several regions of the NOESY spectra, however analysis of the sugar puckers could be performed for the majority of residues. For the native triplex, all nucleotides analysed were found to preferentially adopt the south sugar pucker which is typical for B-type DNA. While it has been reported that the structure of a triplex can be somewhere in between an A- and B-type structure[9], these results suggest that the non-modified triple helix more closely resembles B-DNA. However, sugar puckers could not be determined for six residues within the purine strand of the duplex, and for five residues in the TFO. As it is these two strands which are involved in Hoogsteen bonding to form base triplets, structural analysis of these strands is of particular interest and so further quantitative analysis is required.

Sugar pucker determination for the modified triplex indicated that the sugars corresponding to the residues both attached to, and 3′ to, the modification adopt the north pucker which is more typical of ribose sugars. This is in contrast to the unmodified version for which only DNA-like south sugar puckers were measured,

indicating that upon incorporation of the two 3′-S-phosphorothiolate modifications there is a localised conformational shift around the site of modification.

Analysis of the proton chemical shifts provided an insight into the structural effects of incorporating the two sulfur modifications. Direct electronic effects were observed due to the difference in electronegativity of a sulfur atom compared to oxygen, as well as longer range effects such as changes in ring current shifts caused by an increase in base stacking. Whilst the largest and majority of chemical shift differences are localised around the site of modifications, there are also observed changes along the full length of the oligonucleotide and across all three DNA strands. This implies that in addition to the localised conformational changes, there is a global structural effect as a result of the modifications; this will be examined in Chapter 5.

4.6 Conclusions

Assignment of the native and modified triplex was completed as fully as possible, following the procedure for a B-type DNA oligonucleotide. From comparison of the non-modified triplex assignments to the hairpin assignments detailed in Chapter 2 it was concluded that binding of the third strand has little effect on the overall structure of the hairpin duplex.

Qualitative analysis was performed to some extent on the native and modified triplexes. The sugars corresponding to the W-C hairpin portion of both triplexes and the non-modified Hoogsteen bonded third strand all were found to preferentially adopt the DNA-like south sugar pucker. Upon incorporation of the 3′-S modifications a conformational shift from south to north was observed for both the sugar attached to the modified residue and the residue 3′ to it.

Chapter 5

Structure Generation for the Native and Modified Triplexes

Chapter 5.   Structure Generation for the Native and Modified Triplexes

5.1 Introduction

A full, quantitative structure determination for both the native and 3′-*S*-phosphorothiolate modified triplexes will provide a greater insight to the structural consequences of third strand binding to the hairpin (both natural and modified third strand) than the sugar pucker considerations that were described in Chapter 4. The structure generation process and details of the final structures are described below.

5.2 Calculation of distance restraints

Structures for the native and modified triplexes were generated using both non-exchangeable and exchangeable proton-proton distance restraints. From the structure generation of the target hairpin detailed in Chapter 3 it was concluded that both the 200 ms and build up methods for calculating distance restraints were valid. The 200 ms method was found to produce slightly higher quality structures in terms of the precision and reproducibility, therefore it was decided that distance restraints for both the native and 3′-*S* modified triplexes would be obtained using the NOESY data collected with a 200 ms mixing time and based on the isolated spin-pair approximation.

Non-exchangeable proton-proton distances were calculated from NOESY spectra recorded in D_2O. Labile proton distances were calculated from NOESY spectra acquired at low temperature (2 °C), in a 9:1 H_2O:D_2O solvent mix. Cross peak

volumes were obtained by integrating each nOe signal, and these were converted into distances using the model based on the isolated spin-pair approximation (as described in Section 3.5.1). The C13 H5-H6 cross peak was again used as the reference inter-proton distance. Error boundaries of ± 40 % were applied to all inter-proton distances apart from those corresponding to the terminal residues, which were given looser error boundaries of ± 50 %. The nOe distance restraints files for the native and modified triplex are included in Apendix A2 and A3, respectively.

5.2.1 Additional restraints data

Hydrogen bonding restraints were included in the structure calculations for both the Watson-Crick and Hoogsteen base paired atoms within the native and modified triplexes, using standard hydrogen bonding distances taken from standard values[1] and with applied error boundaries of ± 35 %. Planarity restraints were employed for all base triplets except for those at either end of the molecule.

Loose torsion angle restraints which encompassed typical values for both A- and B-type helices were employed for the sugar phosphate backbone of the Watson-Crick base paired strands of the triplexes. It was anticipated that the 3′-*S* modification would have an effect on the conformation of several sugar puckers in the third strand of the modified triplex, therefore torsion angles for the Hoogsteen bonded strand were not restrained so as not to influence the conformation of the third strand deoxyribose sugars in the energy minimised structures[157].

Table 5.1 gives the number and types of restraints used in the structure generation for both triplexes. The average number of nOe restraints per residue was 12 for both triplexes, which is slightly less than the preferable value of 15. However, for both

sets of structures, both inter- and intra-residue restraints were employed for all nucleotides within the sequence as well as several inter-strand restraints.

Table 5.1. Restraints used in the structure generation for the non-modified and 3′-*S* modified triplexes.

Constraint type	Native triplex	Modified triplex
Non-exchangeable nOe	356	361
Exchangeable nOe	63	51
Hydrogen bond	108	108
Planarity	10	10
Torsion angle	125	125

5.2.2 Simulated annealing protocol

Simulated annealing was performed using the molecular modelling program CNS_solve, and an identical simulated annealing protocol was employed for the triplex structure calculations as was used for structure generation of the target hairpin (see Table 3.2). Standard molecular topology, parameter and link files which include atom masses, atom charges, bond lengths and bond angles were modified (details of which are given in Appendix A4). Two new residues were created for use in the third strand of the triplexes; PCYT – a cytidine nucleotide protonated at the N3 imino position, and STHY – a 3′-*S* modified thymidine residue.
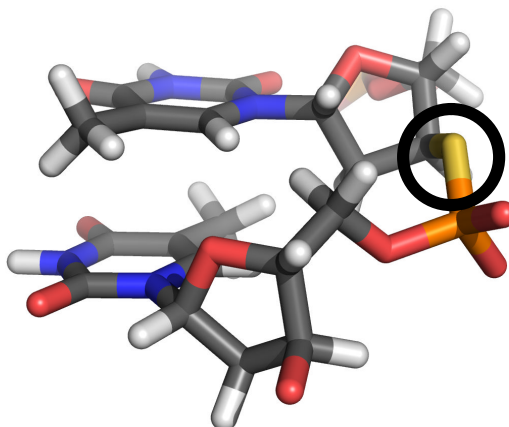
Figure 5.1. Section of the non-modified triplex structure showing one of the modified STHY residues. The sulfur atom is coloured yellow and is circled.

Ten rounds of annealing were performed using torsion angle molecular dynamics with each round using a different initial velocity at which the strands approach one another. Simulated annealing was terminated once 100 acceptable structures had been generated, and the geometry of the 10 lowest energy structures was then saved, and further refined using Cartesian molecular dynamics. A final ensemble of 10 energy minimised structures was generated for each triplex.

5.3 Generated structures

For both the native and modified triplexes, an average structure was calculated from the 10 lowest energy acceptable structures. As can be seen from Figure 5.2 and Figure 5.3, both structures adopted a right-handed helical form. The RMSD from ideal values of bonds and angles for the average of the 10 lowest energy structures were calculated and are given in Table 5.2.

Table 5.2. RMSD from ideal values for the average of the ten lowest energy triplex structures. The RMSD values for the hairpin 200 ms structures are also shown for comparison.

Average Structure	RSMD of bonds / Å	RMSD of angles / °
200 ms hairpin	0.0023	0.4612
Non-modified triplex	0.0018	0.4425
Modified triplex	0.0032	0.7470

The triplex RMSD values were comparable to those obtained for the hairpin structures and were well below the limits for acceptance, which indicates that the structures had good covalent geometry[158]. The non-modified triplex structures had the lowest RMSD values for both bonds and angles, indicating that these structures had the greatest precision.
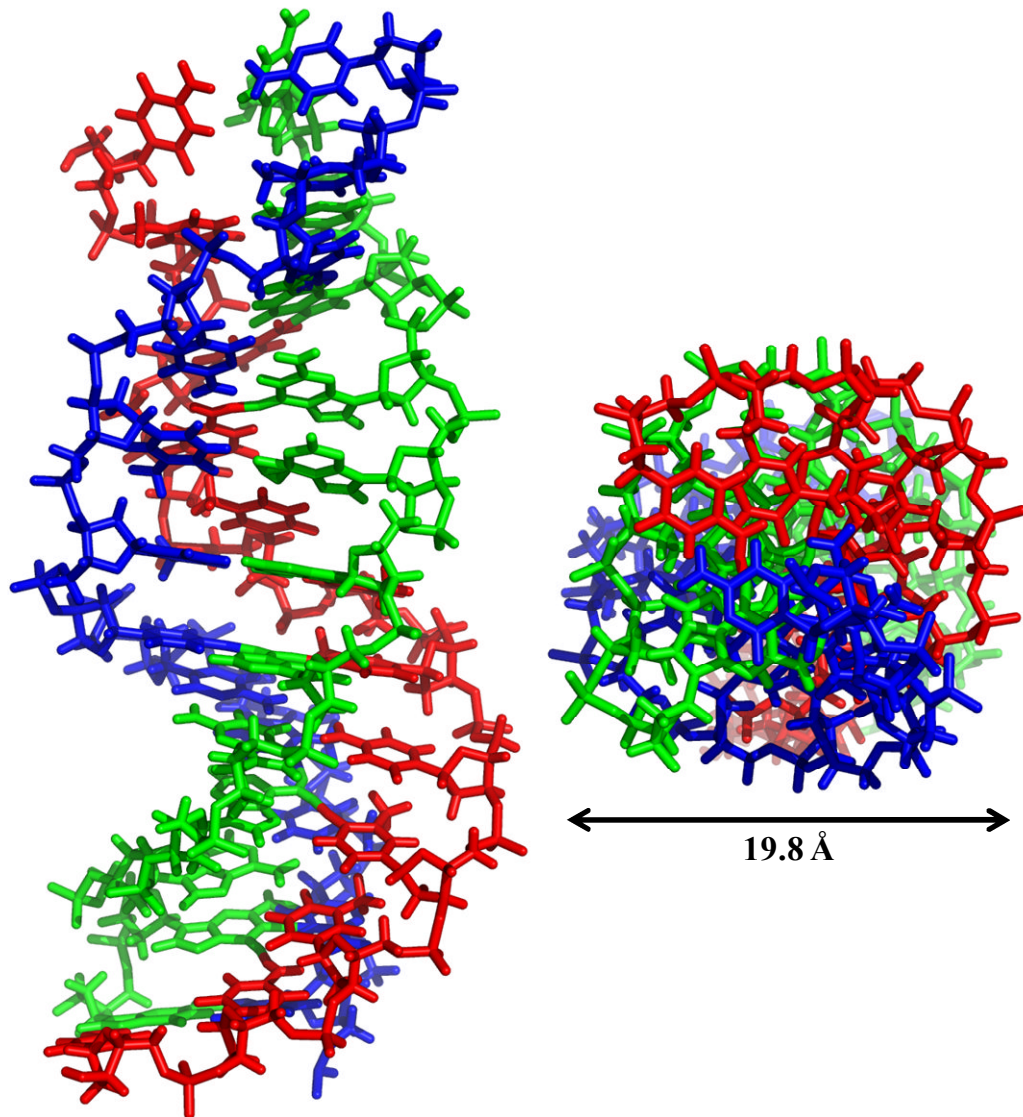
Figure 5.2. Average of the 10 lowest energy structures calculated for the non-modified triplex, with the purine strand coloured green, the Watson-Crick pyrimidine strand blue and the Hoogsteen pyrimidine strand red. The view from the side is shown on the left and view from along the helical axis on the right. The helix diameter was measured at 19.8 Å.
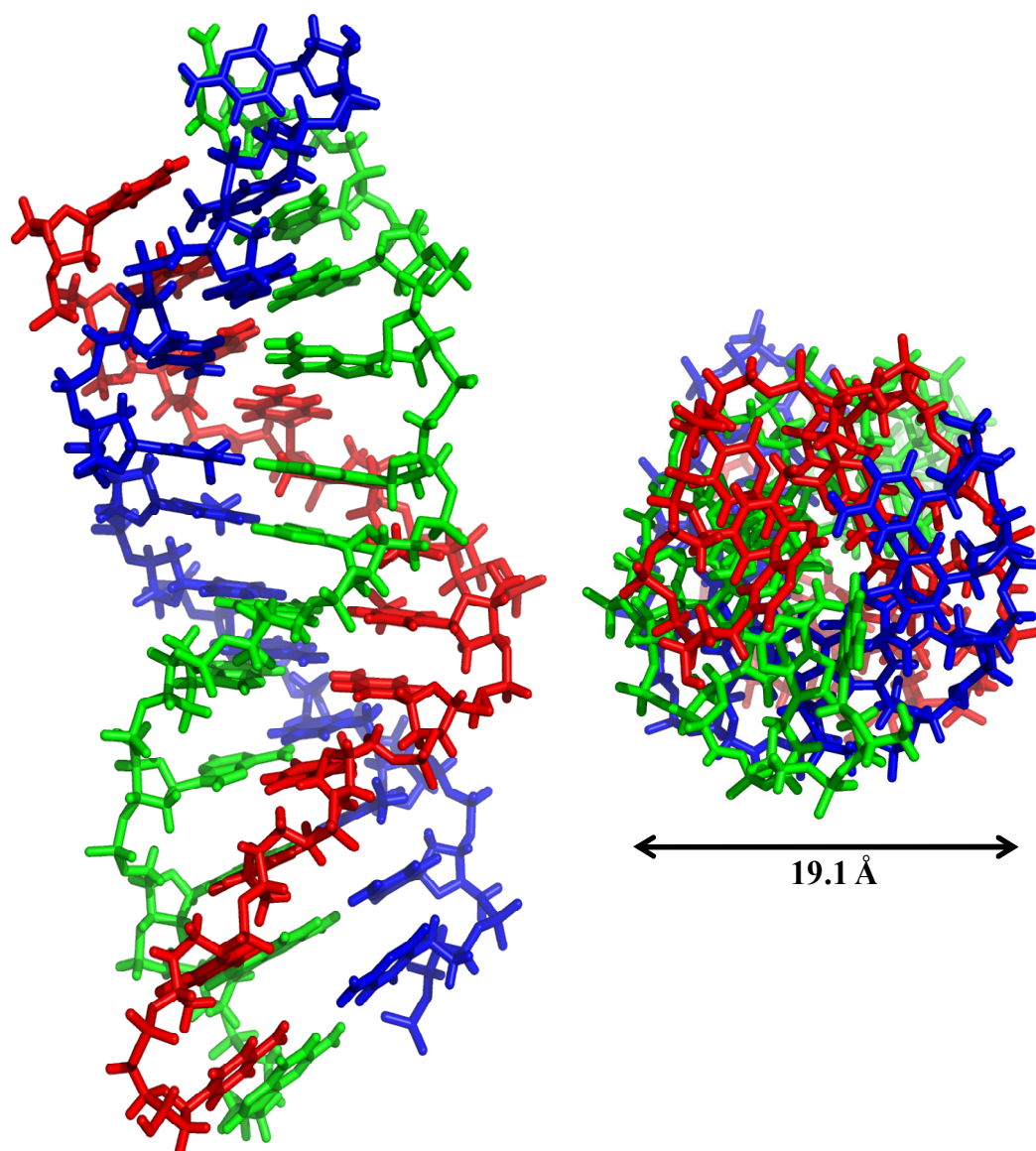
Figure 5.3. Average of the 10 lowest energy structures calculated for the modified triplex, with the purine strand coloured green, the Watson-Crick pyrimidine strand blue and the Hoogsteen pyrimidine strand red. View from the side is shown on the left and view from along the helical axis on the right. The helix diameter was measured at 19.1 Å.

5.3.1 Energy terms

The energy values for the average of the lowest energy triplex structures are presented in Table 5.3. The triplex structure has favourable energies which suggests that under the experimental NMR conditions (pH 5, 20 °C), triplex formation is favourable. The relatively low energies compared to those for the reference B-type duplex suggests that the triplex structure is more extensively defined by experimental constraints than the reference B-type structure, which was generated using 10 restraints per residue.

Table 5.3. Energy values (in kJ mol^{-1}) for the lowest energy triplex structures. Values obtained for the 200 ms hairpin structures and for a B-type DNA duplex[140] are also given for reference.

Energy	Non-modified triplex	Modified triplex	Hairpin (200 ms structure)	B-type duplex[140]
Total	392	555	181	899.0
nOe	60	97	21	35.0
Dihedral	4	7	3	1.1
Bond	23	26	9	16.7
Angle	134	184	61	1751.0
Improper	12	29	9	19.4
Planarity	1	7	1	-
Van der Waals	157	205	78	− 434.0

Both triplex structures have favourable energies compared with the reference data set, however in comparison to the hairpin structure the triplexes have significantly higher energy values. This may be a result of employing fewer nOe restraints per residue (13 restraints per residue for the hairpin and only 11 for the native and modified triplexes).

The modified triplex has higher energy values than the natural equivalent. As a similar number of nOe restraints were employed in both structure calculations, the difference in energies could be a reflection of the type of restraints used, as more exchangeable proton distances were included in the non-modified triplex structure calculations.

5.3.2 Atomic RMSDs

The 10 lowest energy structures that were generated for each set of triplex structures were superimposed and are shown in Figures 5.4 and 5.5. Atomic RMSD values were calculated for the ensemble and are presented in Table 5.4. Each residue within the sequence was found to have an RMSD value significantly lower than 1, indicating that the generated structures have excellent convergence. RMSD values were generally larger at the ends of the molecule which were more loosely restrained.
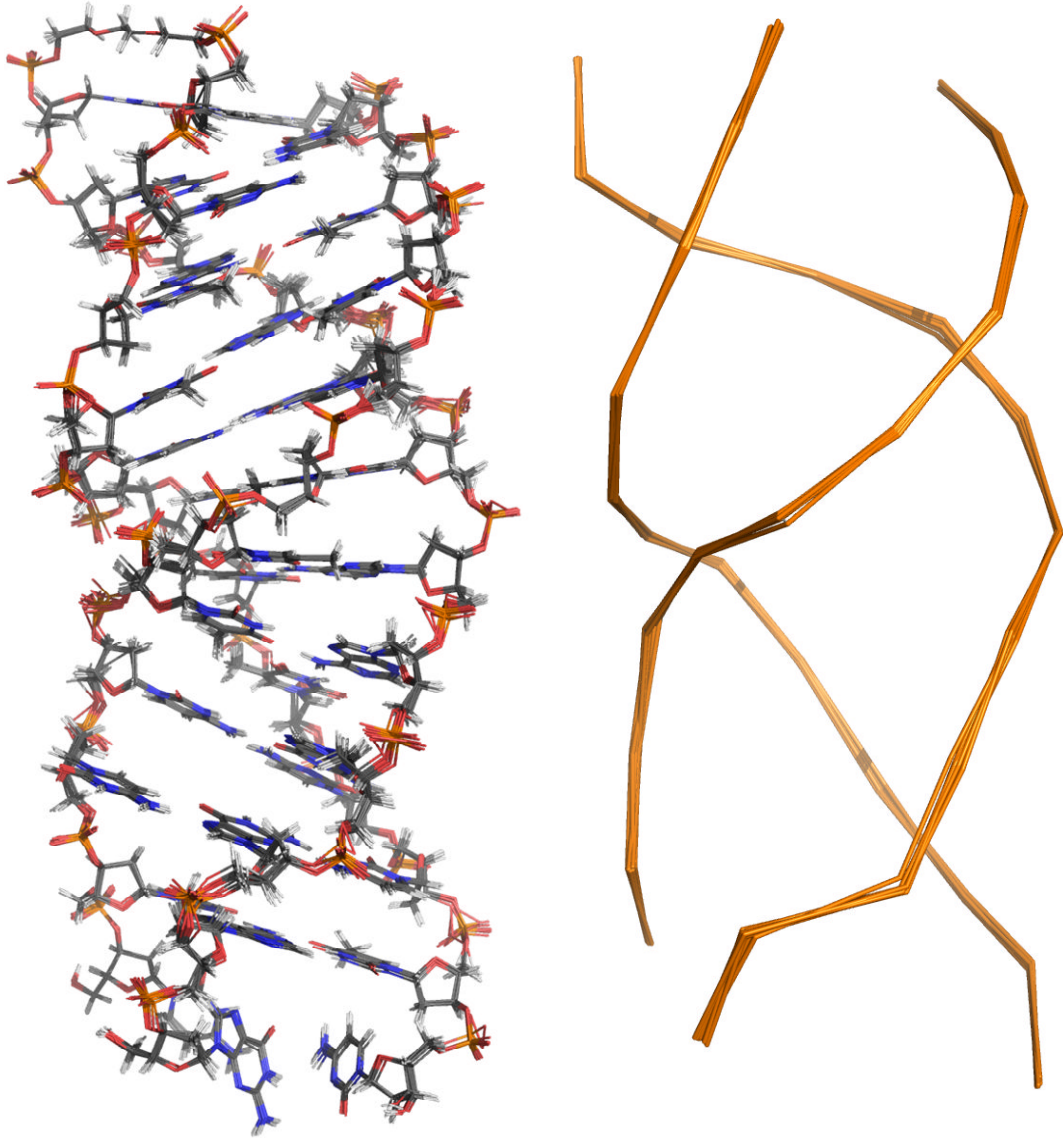
Figure 5.4. Superposition of the 10 lowest energy structures generated for the native triplex. The molecular structure is shown on the left and a ribbon representation of the sugar phosphate backbone on the right.
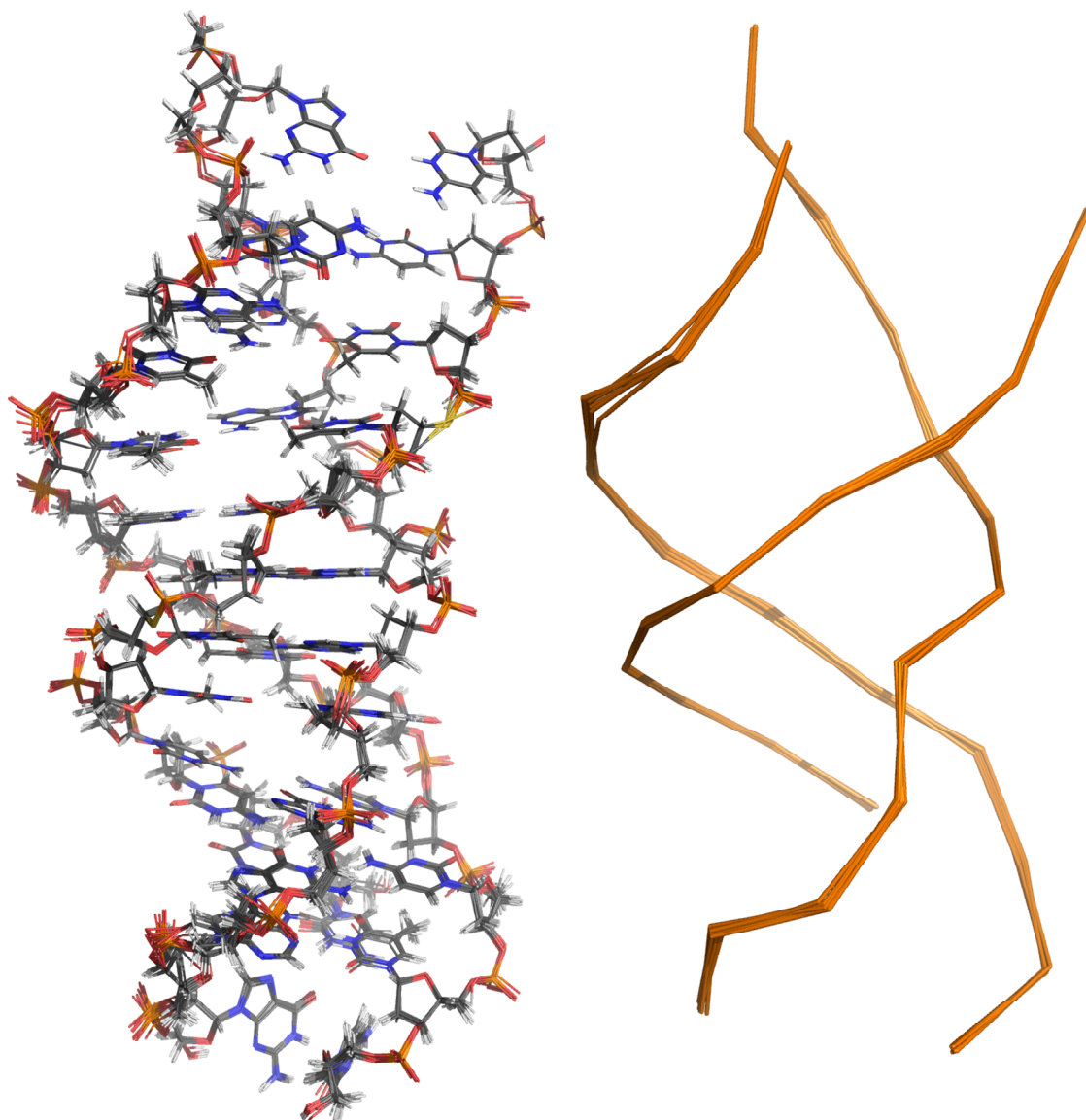
Figure 5.5. Superposition of the 10 lowest energy structures generated for the modified triplex. The molecular structure is shown on the left and a ribbon representation of the sugar phosphate backbone on the right.

Table 5.4. Atomic RMSD values for the ten lowest energy triplex structures, on a per residue basis.

Residue	Native triplex	Modified triplex	Residue	Native triplex	Modified triplex
G1	0.243	0.280	T19	0.140	0.184
A2	0.106	0.217	T20	0.139	0.182
G3	0.159	0.128	C21	0.223	0.167
G4	0.304	0.167	C22	0.178	0.180
A5	0.371	0.160	T23	0.319	0.182
A6	0.150	0.155	C24	0.493	0.198
A7	0.149	0.142	C25	0.506	0.197
G8	0.214	0.147	T26	0.303	0.174
A9	0.356	0.125	C27	0.199	0.129
A10	0.573	0.124	C28	0.122	0.099
G11	0.462	0.143	T29	0.108	0.111
G12	0.562	0.121	**T30/Ts30**	**0.157**	**0.122**
C13	0.477	0.148	T31	0.166	0.194
C14	0.424	0.185	C32	0.251	0.194
T15	0.382	0.226	**T33/Ts33**	**0.331**	**0.177**
T16	0.464	0.274	T34	0.187	0.140
C17	0.580	0.281	C35	0.305	0.198
dU18	0.325	0.288	C36	0.498	0.205
			Average	**0.304**	**0.176**

5.3.3 Sugar puckers

Phosphate-phosphate distances can give an indication of the puckering mode that is adopted by each individual deoxyribose ring. A sugar which resides in a south conformation pushes the adjacent phosphate groups approximately 7.0 Å apart, whereas sugars which adopt a north pucker are associated with shorter inter-phosphate distances of around 5.9 Å[1] (Figure 5.6).
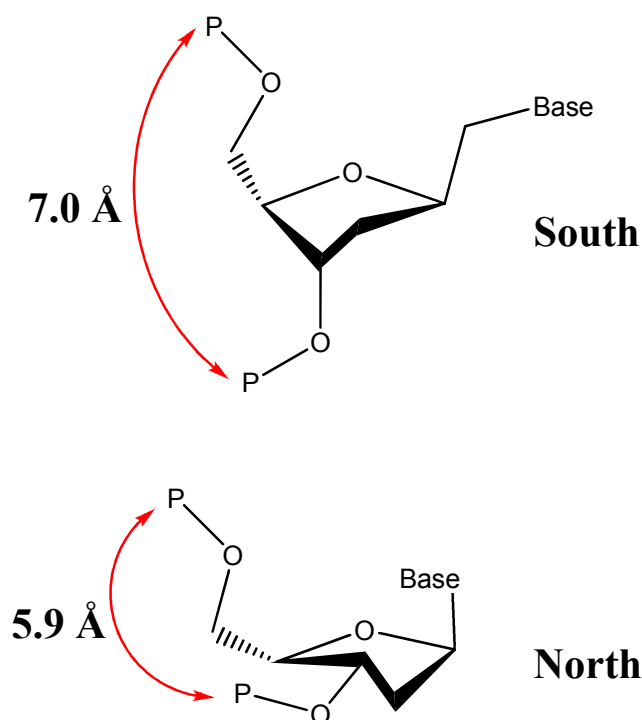
Figure 5.6. Typical phosphate-phosphate distances that are observed in S and N type sugar conformations.

The phosphate-phosphate distances between adjacent phosphorus atoms in the third strands of the native and modified triplex structures were measured and are presented in Table 5.5. Sugar puckers for the residues at the 3′ terminus of the strand could not be determined in this manner due to the lack of a phosphate group at the 3′ end of the oligonucleotides. The inter-phosphate distance between modified residue 6 and residue 7 in the 3′-S third strand is illustrated in Figure 5.7.

Table 5.5. Inter-phosphate distances between residues in the third strand of each of the triplexes (in Å). The corresponding sugar pucker is given in parenthesis.

Residue	Native triplex inter-phosphate distances	Modified triplex inter-phosphate distances
1	**6.1 (N)**	**6.5 (S/N)**
2	6.7 (S)	6.9 (S)
3	7.2 (S)	7.3 (S)
4	6.8 (S)	6.8 (S)
5	6.9 (S)	7.2 (S)
6	6.8 (S)	**5.7 (N)**
7	6.8 (S)	**6.4 (N)**
8	6.7 (S)	6.7 (S)
9	7.0 (S)	**6.1 (N)**
10	6.8 (S)	**6.3 (N)**
11	6.9 (S)	6.6 (S)
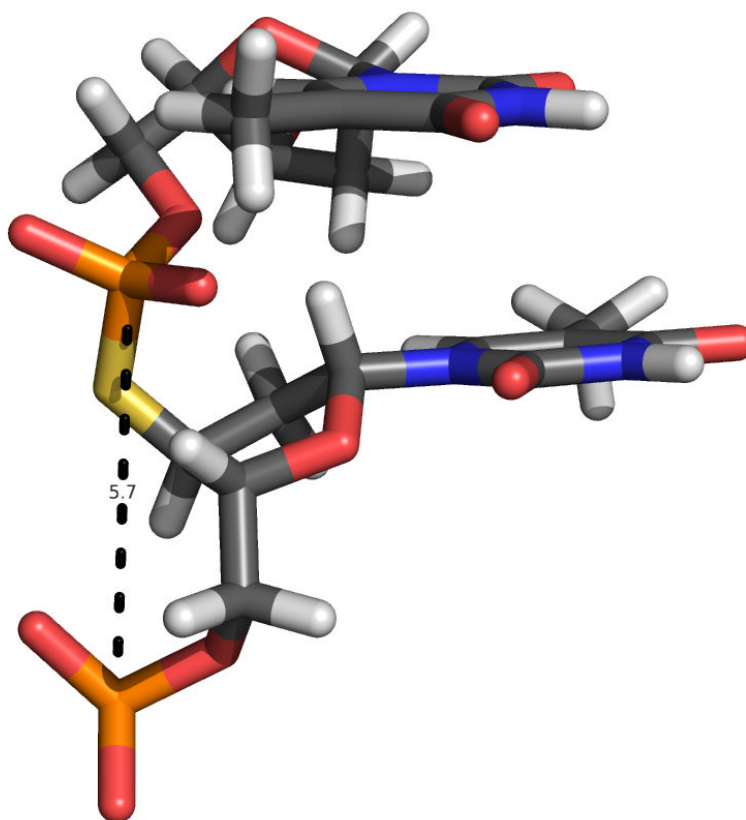
Figure 5.7. Section of the modified triplex structure showing third strand residues 6 and 7. The inter-phosphate distance is labelled (value is in Å).

With the exception of residue 1, all the phosphate-phosphate distances and the corresponding sugar puckering modes are in agreement with those determined using the nOe cross peak heights (detailed in Chapter 4). The relatively short phosphate-phosphate distances of 5.7 Å and 6.1 Å for the two deoxyribose rings that were attached to the sulfur atoms suggests that these sugars are almost exclusively north. The slightly longer distances (6.4 Å and 6.3 Å) that were measured for the sugars 3′ to the modifications imply that these also adopt a north conformation, although to a lesser extent. This is consistent with the conformational changes observed in 3′-*S* modified trinucleotides[92] and DNA:DNA duplexes[94].

5.3.4 Torsion angle analysis

The Watson-Crick hairpin and the Hoogsteen third strand portions of the lowest energy structures were analysed separately using CURVES[+141]. Torsion angles along the phosphate backbone for both parts of the native and 3′-*S* triplexes were obtained and average values are displayed in Table 5.6 and Table 5.7.

The torsion angles along the backbone of the hairpin section of both triplexes (Table 5.6) are comparable to those along the backbone of the target hairpin, and both are consistent with typical values for a B-type helix. There are only slight variations in several of the backbone torsion angles between the native and modified triplexes, indicating that the global structure is altered only slightly upon binding of the modified third strand, compared to binding of the natural equivalent.

Table 5.6. Average backbone torsion angles (in degrees) for the Watson-Crick hairpin section of the triplex structures. Torsion angles calculated for the 200 ms hairpin structure as well as typical values for standard B-type[142] and A-type[143] DNA are also given for comparison.

Structure	α	β	γ	δ	ε	ζ	χ
W-C portion of native triplex	− 58	− 145	44	130	− 143	− 90	− 102
W-C portion of modified triplex	− 53	− 144	53	134	− 144	− 91	− 83
Hairpin 200 ms	− 49	− 138	40	141	− 140	− 95	− 69
B-type DNA	− 46	− 147	36	157	− 133	− 96	− 98
A-type DNA	− 75	− 179	59	79	− 155	− 67	− 158

The torsion angles along the phosphate backbone of the Hoogsteen bonded native and modified third strands (Table 5.7) are somewhat intermediate between the typical values for A- and B-type DNA, with the modified version being more similar to A-DNA than the native equivalent. There is a significant difference in torsion angles between the unmodified and modified versions which suggests a significant change in structure. This is consistent with the conformational shift of several deoxyribose rings in the modified third strand.

Table 5.7. Average backbone torsion angles (in degrees) for the Hoogsteen bonded third strand of the triplex structures. Typical values for standard B-type[142] and A-type[143] duplexes are also given.

Structure	α	β	γ	δ	ε	ζ	χ
Third strand of native triplex	− 58	− 145	44	130	− 143	− 90	− 102
Third strand of modified triplex	− 74	− 155	50	131	− 140	− 79	− 124
B-type DNA	− 46	− 147	36	157	− 133	− 96	− 98
A-type DNA	− 75	− 179	59	79	− 155	− 67	− 158

5.3.5 Helical parameter analysis

The helix diameters were measured as inter-strand phosphate-phosphate distances between the two W-C bonded strands, at several different points along the length of the helix; this is illustrated in Figure 5.8. The average values for the native and modified triplexes were calculated at 19.8 Å and 19.1 Å respectively, as indicated in Figures 5.2 and 5.3. Both are consistent with the average helical diameter of B-type DNA, which is 20 Å. The major groove width of the W-C hairpin portion of the triplex structures were measured and are given in Table 5.8. Values were found to be comparable to B-type helices for both the non-modified and modified structures.
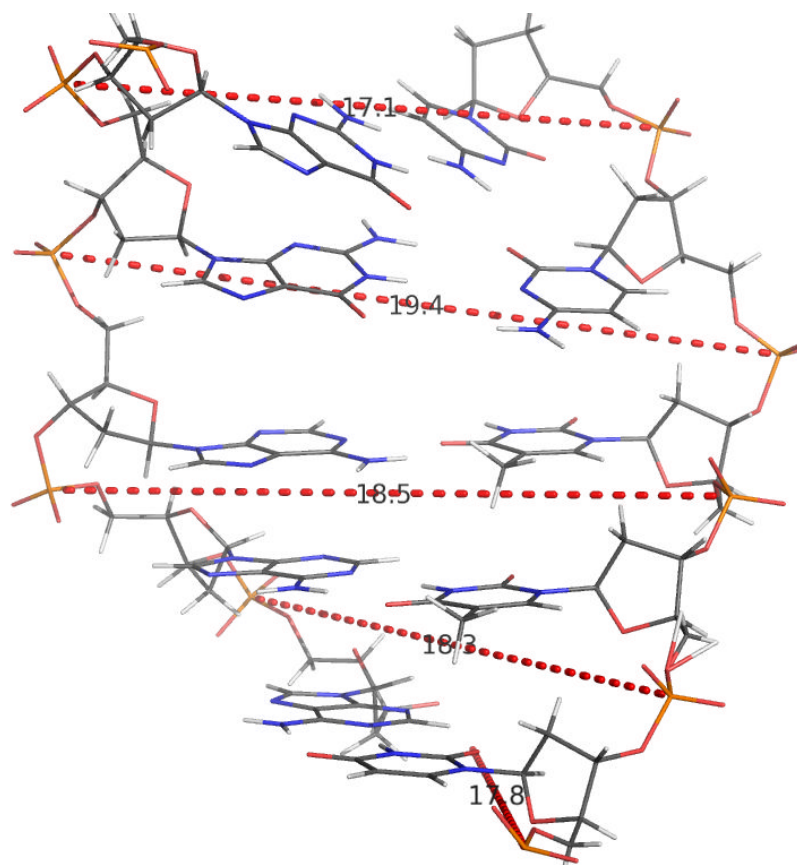
Figure 5.8. Section of the modified triplex structure showing several helix diameter measurements.

Table 5.8. Average major groove widths for the Watson-Crick portion of the triplex structures. Values were calculated as refined P-P distances, taking into account the directions of the sugar-phosphate backbones[145]. The calculated hairpin 200 ms structure, and typical values for B-DNA[146] and A-DNA[147] are given for reference.

Structure	Major groove width /Å
W-C portion of native triplex	17.8
W-C portion of modified triplex	18.2
Hairpin 200 ms	16.8
B-DNA	17.3
A-DNA	13.6

Helical analysis of the W-C duplex section of the triplex structures were obtained using CURVES[+141] and average values for the base step parameters are presented in Table 5.9. Base pair step parameters for both the native and modified triplexes are generally comparable to the hairpin with a slightly reduced tilting of the bases, particularly for the modified triplex.

The base pairs within the triplex structures are more tilted than in a standard B-type duplex, however the separation between base pairs (base pair rise) is close to ideal. Base pairs within the native triplex are generally less twisted than in a B-type duplex, although this is not true for the modified version. Base pair shift and slide values are both comparable to the shift and slide values within B-DNA.

Table 5.9. Average base pair step parameters for the lowest energy triplex structures. Values obtained for the hairpin 200 ms structure and for an idealised B-type helix are given for comparison[146].

Base pair step parameter	W-C portion of native triplex	W-C portion of modified triplex	Hairpin (200 ms)	Ideal B-type helix
Tilt (°)	3.90	2.85	4.71	-0.19
Roll (°)	9.62	10.10	8.10	0.02
Rise (Å)	3.49	3.75	3.94	3.36
Twist (°)	23.39	33.73	35.33	35.58
Slide (Å)	0.78	0.15	0.86	0.12
Shift (Å)	0.84	1.44	0.92	-0.02

Figures 5.9 and 5.10 highlight the local helical changes along the sequence for the W-C portion of the unmodified triplex. As with the hairpin, there are much greater fluctuations in the local structure of the triplex, compared to a standard B-type DNA duplex.

Base Pair Tilt (in Degrees)

Base Pair Roll (in Degrees)

Base Pair Rise (in Angstroms)
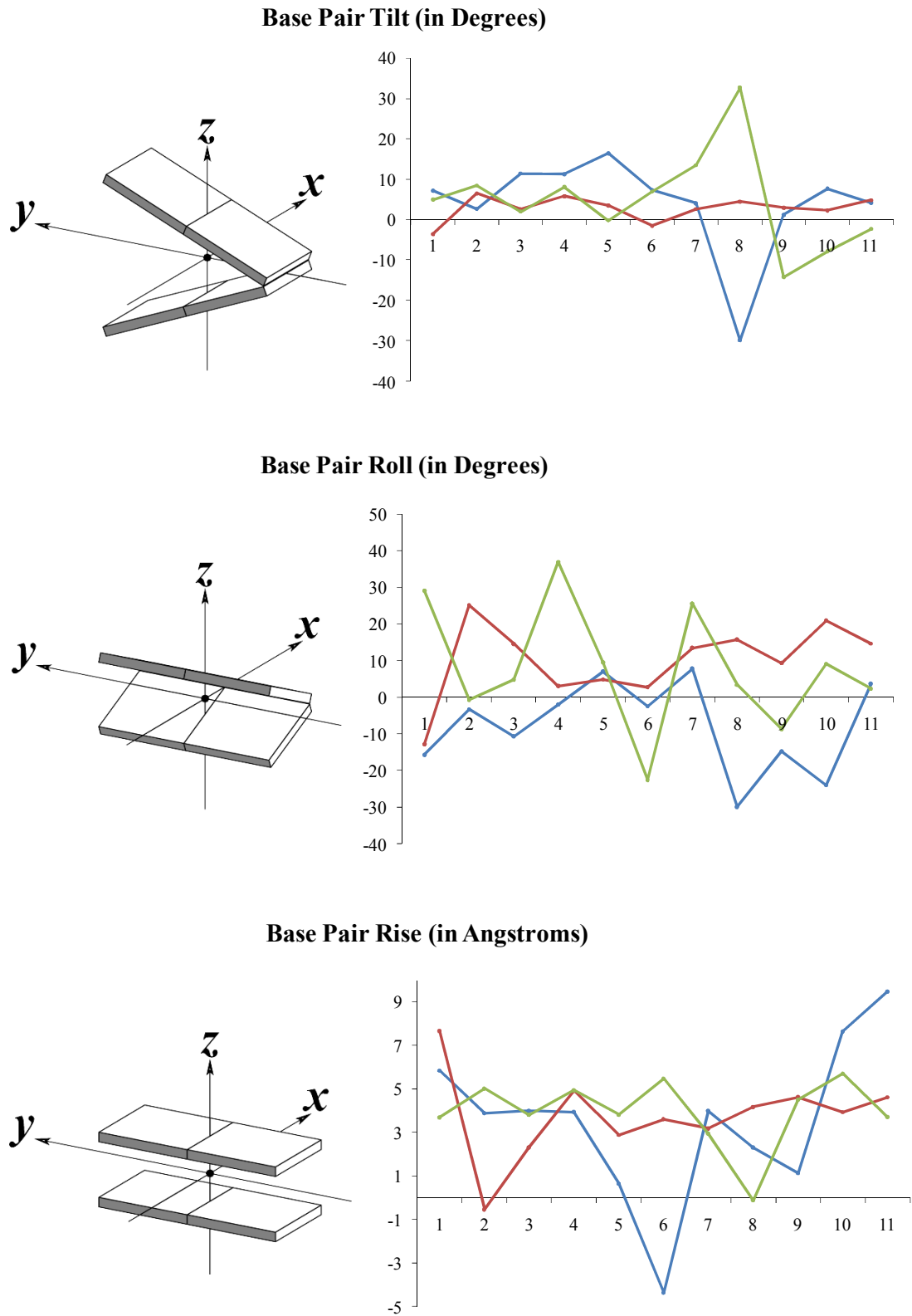
Figure 5.9. Base pair tilt, roll and rise values for the minimum energy native triplex (blue), modified triplex (red) and hairpin 200 ms (green) structures. Numbers along the x-axis indicate base pair steps. Representations of positive values of each designated parameter[148] are also illustrated on the left.

Base Pair Twist (in Degrees)

Base Pair Slide (in Angstroms)

Base Pair Shift (in Angstroms)
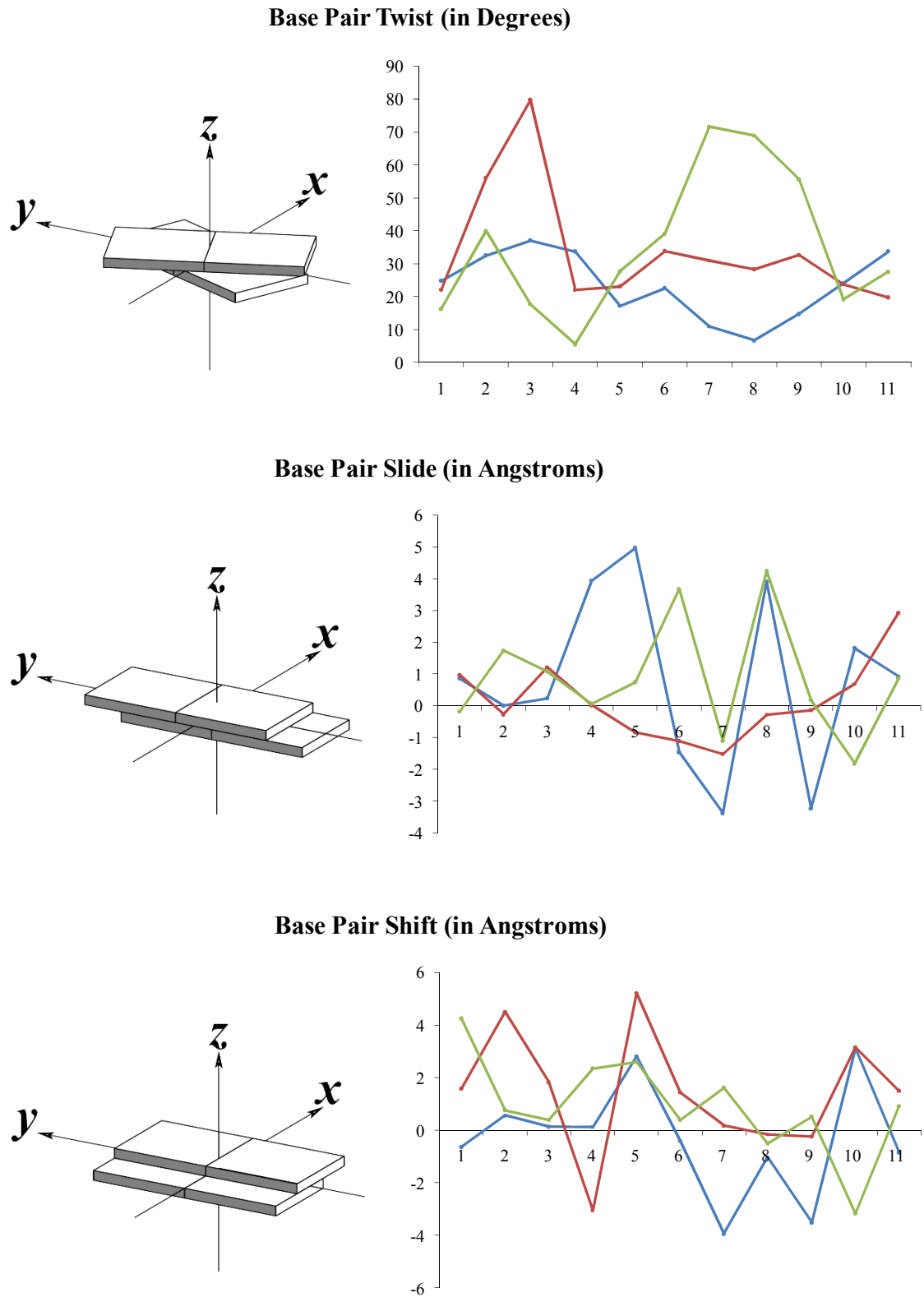
Figure 5.10. Base pair twist, slide and shift values for minimum energy native triplex (blue), modified triplex (red) and hairpin 200 ms (green) structures. Numbers along the x-axis indicate base pair steps. Representations of positive values of each designated parameter[148] are also illustrated on the left.

5.4 Discussion

The native and modified triplex structures generated both have favourable energy values and no restraint violations, indicating that the experimental restraints used in structure calculations were sufficient. There is a significant difference between the energies for the native and modified triplex structures, with the modified having a total energy of over 100 kJ mol^{-1} higher than the natural equivalent. As a similar number of nOe restraints were employed in each structure calculation, the difference in energy may have resulted from differences in the type of restraints used, as more restraints involving exchangeable protons were included in the unmodified triplex structure calculation. The exchangeable set of nOe restraints include distances between atoms which are involved in hydrogen bonding, as well as inter-strand distances, which suggests that this information is important in order to determine higher quality structures.

Likewise, the modified triplex structures had higher RMSD from ideal values than the non-modified version (although still well below the limits for acceptance), showing a loss in precision; again possibly as a result of the number of each type of nOe restraints.

The puckering modes adopted by individual sugar rings that were determined by measuring inter-phosphate distances are mostly in agreement with the sugar pucker analysis detailed in Section 4.4.2, with some discrepancies at the ends of the molecules. As the terminal residues were only loosely restrained the structure is likely to be less accurate at the ends which may lead to greater errors in the analysis of these residues. The phosphate-phosphate distances calculated were generally intermediate between the distances given in the literature for the two conformational

extremes, which highlights the position of equilibrium between the two puckering modes for each individual sugar ring.

The triplex was separated into two components for backbone torsion angle analysis; the Watson-Crick bonded duplex and the Hoogsteen third strand. The backbone torsion angles within the duplex do not alter much upon incorporation of the third strand, which implies that binding of the third strand in the major groove causes minimal disruption to the structure of the hairpin. This is in agreement with both the chemical shifts presented in Chapter 4 and previous literature[156]. Torsion angles along the backbone of the Watson-Crick portion of the native and modified triplexes were fairly similar, therefore the chemical shift differences which were presented in Section 4.4.1 must be a result of localised variations, which do not have a significant effect on the overall structure of the molecule.

A greater change was observed in the third strand torsion angles upon inclusion of the modifications. In the unmodified third strand, the majority of torsion angles were comparable to those observed in B-DNA, however the modified third strand was more similar to an A-type helix. This implies that as anticipated, the third strand becomes more RNA-like as a result of the 3'-S-phosphorothiolate modifications.

The helical parameters within the duplex portion of the triplexes were found to be comparable to those measured for the target hairpin structures, with only slight differences as a result of accommodating the third strand. In agreement with the analysis detailed in Chapter 3, greater variations in base pair step parameters along the length of the sequence were observed compared to a standard B-type DNA dodecamer, which reiterates the importance of using a suffecient number of experimental constraints in the structure calculations.

5.5 Conclusions

High quality, low energy structures were generated for the native and triplex structures using nOe distance constraints that were calculated using the 200 ms method. All structures produced had favourable energies, no violating constraints and low RMSD values.

The structure of the target hairpin was essentially unaltered upon binding of either the natural or the modified TFO, and retained a B-type helix. The third strand became more RNA-like upon inclusion of the 3′-*S*-phosphorothiolate linkers by altering the conformation of the sugar rings to which they were attached, and the ring 3′ to the modification.

Chapter 6

UV Thermal Analysis

Chapter 6. UV Thermal Analysis

6.1 Introduction

A knowledge of triplex thermodynamics is imperative given the potential biological and therapeutic applications of DNA triplexes. In order to design and develop triplex molecules for use as therapeutic agents it is essential to have an understanding of factors which affect triplex stability, such as the sequence, base composition and solution conditions. The concept of UV thermal analysis of nucleic acid structures was introduced in Chapter 1; this chapter describes the UV analysis of the hairpin, native triplex and modified triplex that was carried out in order to determine the stability of each structure and the thermodynamic consequences of incorporating two 3′-S-phosphorothiolate modifications into the third strand of a parallel DNA triplex.

6.2 UV melting profiles of triplex DNA

Two melting transitions occur during UV thermal analysis of triplex structures, which can be assigned to (1) the dissociation of the third strand, leaving a duplex, and (2) the duplex to single strand[159]. Depending on solution conditions, the two transitions can occur simultaneously producing a monophasic curve, or a biphasic curve may be apparent due to the initial dissociation of the Hoogsteen base-paired strand, followed by duplex dissociation[159, 160] (Figure 6.1).
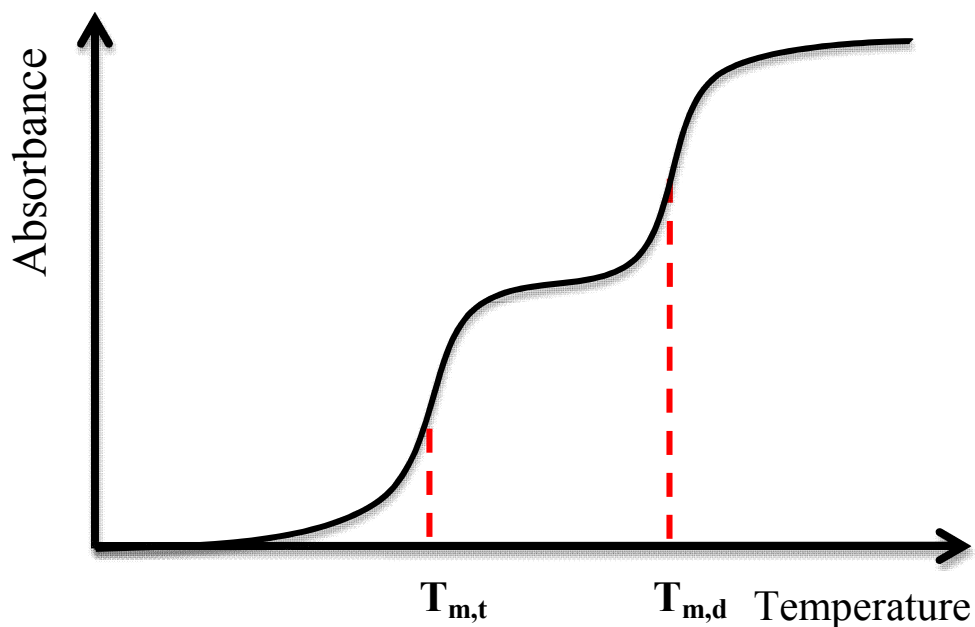
Figure 6.1. Schematic of a biphasic melting curve which is characteristic of some DNA triplexes. The triplex to duplex and duplex to single strand melting transitions are labelled as $T_{m,t}$ and $T_{m,d}$, respectively.

Triplexes can be classified into two categories; antiparallel and parallel, based on the third strand composition and orientation, as previously described in Chapter 1. Both classes of triplex require different solution conditions for their formation, which renders comparison of different triplex structures by UV thermal analysis problematic as data would need to be collected under identical solution conditions.

Relatively little information is available regarding the thermal stability of antiparallel triplexes, which are difficult to monitor by UV melting curves due to the fact that only a very small hyperchromism is observed for these structures[20]. However one example which utilised circular dichromism techniques reported that an antiparallel triplex composed of GAA repeats in the third strand was stable at physiological pH and room temperature[161].

In contrast, there is much literature precedence on the stability of parallel triplexes. The composition of the pyrimidine third strand affects the melting behaviour of the triplex; those which contain only T:A-T base triplets (i.e. no cytosine residues in the third strand) melt in monophasic transitions under all solution conditions whereas those containing C^+:G-C triplets can, depending on the solution conditions, exhibit a biphasic melting curve from which two melting temperatures (T_m) may be obtained[160].

The stability and melting behaviour of triplexes which contain cytidine residues in the third strand is greatly affected by the pH. A parallel intramolecular triplex comprising three thymidine and three cytidine residues in the third strand was reported by Soto *et al*. to exhibit monophasic melting behaviour at acidic pH, whereas at physiological pH a biphasic curve was observed[160]. This indicates that the triplex is most stable under acidic conditions as the triplex to duplex unfolding occurs at an increased temperature compared to that at neutral pH, causing it to overlap with the duplex to single strand melting transition. This is consistent with the requirement for protonation of third strand cytosine residues for triplex formation[41, 159, 162]. At low pH, triplexes rich in C^+:G-C triplets are more stable than those containing exclusively T:A-T triplets, due to the formation of stronger Hoogsteen hydrogen bonds that are associated with C^+:G-C triplets, along with favourable electrostatic interactions between the positively charged cytosine base and the negatively charged phosphate group in the backbone[151, 153].

The effect of ionic strength on triplex stability has been investigated using UV thermal analysis and differential scanning calorimetry (DSC) techniques[160]. Several duplexes and triplexes containing only thymidine residues in the third strand were found to be more stable upon increasing sodium concentration, due to stabilisation of

the negatively charged phosphate backbones[160]. In contrast, increasing salt concentration had a destabilising effect on triplexes which contain third strand cytidine residues. The presence of Na^+ cations around the positively charged cytosine bases results in a decrease in the pK_a of the triplex, and so at higher salt concentrations a lower pH is required to induce cytosine protonation[160].

The positioning of cytidine residues in the third strand has also been found to affect the stability of a parallel triplex. Several studies have reported that adjacent cytidines are destabilising to the structure[49, 159], which has been attributed to electrostatic repulsion between positive charges on neighbouring cytosine bases. In addition, a smaller base stacking energy is associated with C_n sequences compared with $(CT)_n$ sequences[153, 163]. Unfavourable electrostatic interactions only occur between protonated cytosines, so this effect is most evident as the pH is decreased[159], however at neutral pH, deprotonation of the third strand cytosines itself becomes the cause for destabilisation.

The unfolding thermodynamics of DNA triplexes with different molecularities has also been widely studied, and it is recognised that triplex stability follows the order unimolecular > bimolecular > trimolecular due to a lower entropy penalty associated with lower molecularity. The inclusion of connecting loops between two strands can also have a stabilising effect on triplex formation, due to improving base stacking and shielding of external base triplets from the solvent[162]. Triplex molecularity also has an effect on the protonation of cytosine bases in the third as reported by Lee *et al.*[162], who found that the percentage of protonation decreased with increasing molecularity.

In the case of a bimolecular triplex, the stability of the underlying duplex has an impact on triplex formation. A recent study by Rusling *et al.*[164] demonstrated that the

length and sequence of a duplex affects the thermal stability of the triplex. They analysed the interaction of a 14mer TFO with a variety of different duplex contexts which all contained the same 14 base-pair target sequence, but varied in length, sequence and molecularity. They found that increasing the length of the duplex by flanking the sequence by 6 AT or 6 GC base pairs on either side increased the T_m, even though the target site was unchanged. In addition, a larger T_m increase was observed when the target sequence was flanked by GC base pairs than the duplex which contained AT tails. It is well known that duplexes rich in GC base pairs are more stable than those which contain predominantly AT pairs due to an additional Watson-Crick hydrogen bond in GC base pairs, therefore increasing the stability of the underlying duplex increases the stability of the triplex. Finally, they observed an increase in the triplex T_m upon incorporating a hexaethylene glycol linker to connect the two strands of the duplex, which is consistent with the effect of molecularity and inclusion of linkers on triplex formation.

6.2.1 Modifications which improve triplex stability

As described in Chapter 1, many studies have involved using modifications to increase the stability of a parallel triplex such as the replacement of cytosines in the third strand with 5-methylcytosine. This was found to increase both the thermal stability and the upper pH limit for triplex stability due to the increased pK_a value of the electron donating methyl group and increased stacking energies and hydrophobic effects[41, 159].

Replacing the phosphodiester group in the backbone of the third strand of a triplex with non-ionic linkages has also been found to improve the binding affinity, with the

best results associated with more bulky, hydrophobic substituents[165]. Triplex formation has even been observed at neutral pH using such linkers although below the required levels to be potentially used *in vivo*.

The thermodynamic effects of using locked nucleic acids (LNAs) in triplex forming oligonucleotides have been widely studied[166, 167]. Partially modified TFOs have been found to enhance triplex stability by up to 26 °C and enable triplex formation to some extent at pH 7[167].

6.3 Stability of nucleic acids containing 3′-*S* modifications

As discussed in Chapter 1, 3′-*S*-phosphorothiolate modifications have previously been incorporated into nucleic acid structures, with the effect of altering the conformational equilibrium of the modified sugar and to a lesser extent, the 3′-neighbouring sugar from the DNA-like south, to an RNA-like north conformation. Several studies have reported that replacement of the 3′ oxygen for a phosphorothiolate linker within a DNA sequence results in an enhanced binding affinity of the molecule with its complementary RNA strand[94, 96, 168]. An increase in T_m of 2.5 °C was reported for a dodecamer duplex containing a single modification compared with the unmodified equivalent. Two adjacent modifications were found to increase the stability of the structure by 3 °C and two 3′-*S*-phosphorothiolate linkages in non-adjacent positions enhanced the stability even further (ΔT_m + 4 °C). Inclusion of the sulfur modification into an all DNA duplex resulted in destabilisation of the structure, due to the decreased stability of a DNA:RNA hybrid duplex compared to an all DNA or RNA equivalent[96].

The 3′-*S*-phosphorothiolate linkage has also been incorporated into one or more cytidine residues within an i-motif structure, d(TCCCCC)[93]. I-motif sugars preferentially adopt a north conformation; and inclusion of the 3′-*S*-phosphorothiolate modification mimics this without the destabilising bulky OH substituent at the C2′ position[93]. A similar T_m increase of 2.5 °C was observed for the i-motif containing a single modification, however a much higher increase of up to 8.5 °C was observed when two adjacent modifications were included[93].

6.4   UV thermal analysis of the target hairpin

The stability and melting behaviour of the target hairpin over a range of pH values was first analysed in order to enable comparisons to be made between the duplex and triplex. A range of pH values were chosen which encompassed both physiological and slightly acidic pH, a requirement for parallel triplex formation. The UV absorbance of the hairpin at pH 5, 6 and 7 was therefore measured against increasing temperature, producing sigmoidal melting curves typical of the unfolding of DNA secondary structure. The normalised melting profiles are shown in Figure 6.2.
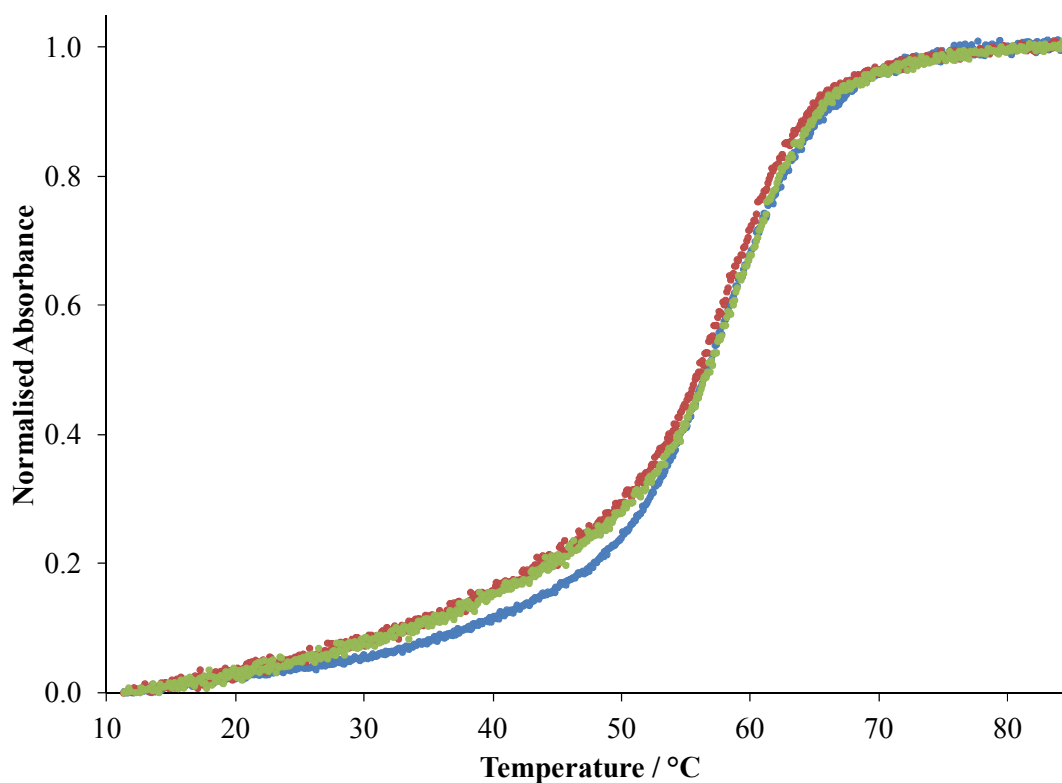
Figure 6.2. Normalised melting profiles for the hairpin at pH 5 (blue), pH 6 (red) and pH 7 (green). Absorbance was measured at 260 nm and samples were heated from 10 - 85 °C at a rate of 1 °C /min. Sample concentrations were 10 µM in 20 mM sodium phosphate buffer.

Melting temperatures were calculated from the first derivative plot of each melting profile and the T_m for each pH value averaged over two runs are given in Table 6.1. The melting temperature of the hairpin at each pH value is around 58 °C, which signifies that the structure predominantly exists in its folded state at physiological temperature (37 °C). The T_m remained fairly constant as the pH was lowered, indicating that the stability is pH independent. This result is consistent with previous studies which investigated the effects of pH on DNA duplex stability[169, 170].

Table 6.1. Melting temperature of hairpin over a range of pH values, averaged over two runs.

pH	T_m / °C
5	57.8
6	57.6
7	57.7

The melting transitions occur over a fairly wide temperature range (approximately 30 °C), which suggests that the Gibbs free energy of the hairpin unfolding reaction has a low temperature dependence[171]. Various thermodynamic parameters can be extracted from the shape of the melting curves, assuming that the system exists in a two state equilibrium (folded and unfolded) and that melting is independent of strand concentration. These assumptions can be applied to the hairpin, due to the monophasic nature of the melting curve and the unimolecularity of the system.

The absorbance vs. temperature plot for the pH 7 hairpin was converted into a fraction folded, θ, vs. temperature plot (Figure 6.3) using Equation 6.1, where A_T represents the absorbance at a given temperature, θ_T, and $L0_T$ and $L1_T$ correspond to the baseline values of the unfolded and folded species, respectively.

$$\theta_T = (L0_T - A_T)/(L0_T - L1_T)$$

Equation 6.1
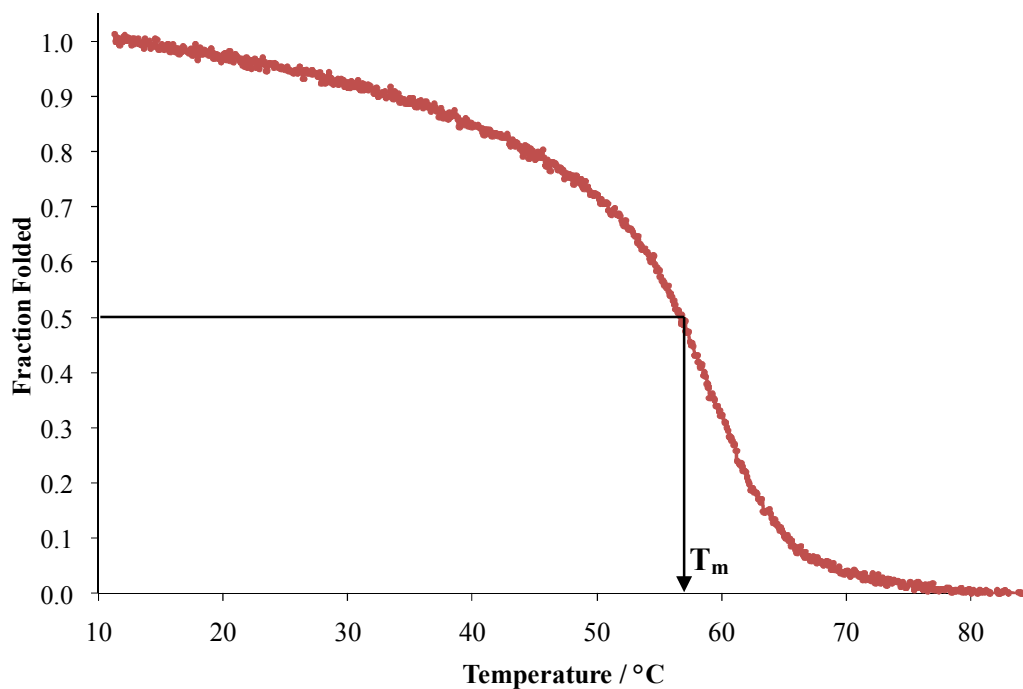
Figure 6.3. Fraction folded, θ as a function of temperature for the hairpin (pH 7).

As mentioned in Chapter 1, the melting temperature is defined as the temperature at which exactly half of the molecules are in an associated, or folded state, which corresponds to θ = 0.5. The melting temperature can therefore be extrapolated from the fraction folded vs. temperature curve, and is indicated on Figure 6.3. The T_m obtained using this method was found to be 56.7 °C, which is slightly lower than the T_m calculated from the first derivative of the melting curve.

The affinity constant, K_a is given by Equation 6.2, and a plot of the natural logarithm of K_a ($\ln(K_a)$) vs. reciprocal temperature ($1/T$, K^{-1}) (Figure 6.4) allows additional thermodynamic information to be obtained, such as the standard van't Hoff enthalpy, $\Delta H°$, standard entropy, $\Delta S°$ and Gibbs free energy, $\Delta G°$.

$$K_a = \theta/(1-\theta)$$ **Equation 6.2**

Analysis was restricted to the temperature range for which $0.15 < \theta < 0.85$, in accordance with literature[105]. Beyond this range almost all of the molecules are fully associated or dissociated and it becomes relatively difficult to evaluate the affinity constant.
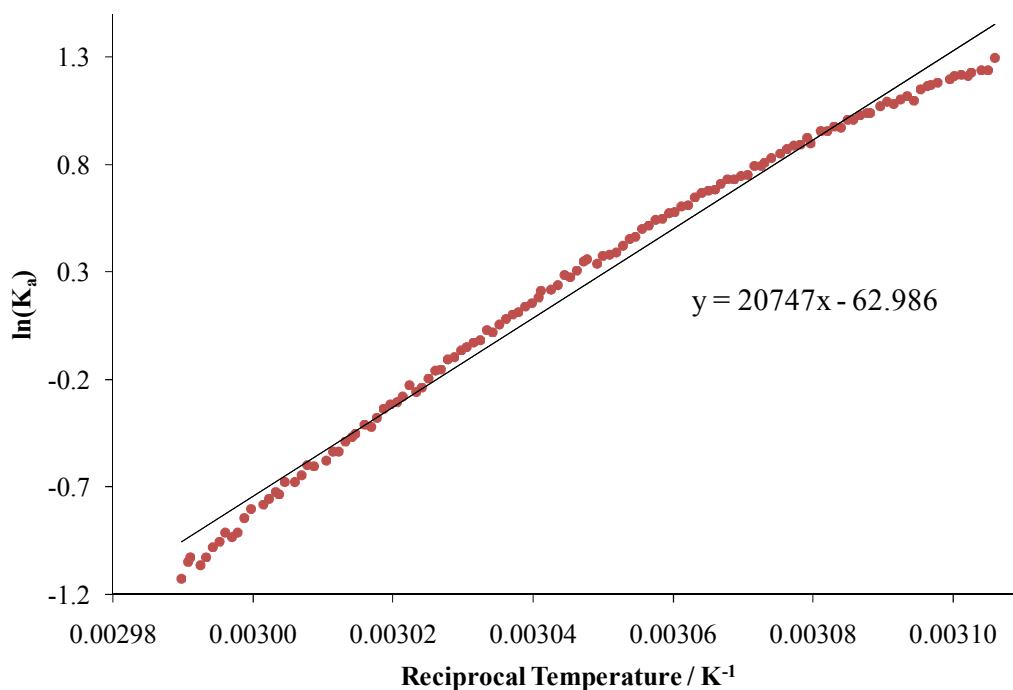
Figure 6.4. Plot of ln(Ka) vs. 1/T for the hairpin (pH 7).

Gibbs free energy is given by Equation 6.3 (where R denotes the Gas constant), which can be rearranged to give Equation 6.4 in the form $y = mx + c$. From the $\ln(K_a)$ vs. 1/T plot we can deduce that the gradient of the linear fit is equal to $-\Delta H^\circ/R$ and the intercept is equal to $\Delta S^\circ/R$.

$$\Delta G^\circ = -RT\ln(K_a) = \Delta H^\circ - T\Delta S^\circ \qquad \textbf{Equation 6.3}$$

$$\ln K_a = -\Delta H^\circ/R \times (1/T) + \Delta S^\circ/R \qquad \textbf{Equation 6.4}$$

These values can then be substituted back into Equation 6.3 to calculate $\Delta G°$ at physiological pH (37 °C, 310 K). The resulting thermodynamic parameters for the hairpin at pH 7 are given in Table 6.2.

Table 6.2. Thermodynamic parameters extracted from the $\ln(K_a)$ vs. 1/T plot for the hairpin at pH 7, 6 and 5. $\Delta G°$ is given at 310 K. Errors are within 10 %.

Thermodynamic parameter	pH 7	pH 6	pH 5
$\Delta H° / kJ\ mol^{-1}$	−172	−186	−171
$\Delta S° / J\ K\ mol^{-1}$	−523	−557	−515
$\Delta G° / kJ\ mol^{-1}$	−10	−13	−11

The large negative standard enthalpies calculated indicate that hairpin folding is an enthalpically driven process. At physiological temperature, the small entropy contributions result in a favourable, negative $\Delta G°$ meaning that folding of the structure is favourable under these conditions.

6.5 UV thermal analysis of the triplex as a function of pH

UV melting profiles were obtained for the non-modified and modified triplex over the same pH range as for the hairpin, with additional runs at pH 6.5. Melting temperatures averaged over two runs are given in Table 6.3.

Table 6.3. Melting temperatures for the UV unfolding of the non-modified and modified triplexes at 260 nm. Values are given in °C.

pH	Non-modified triplex T_m	Modified triplex T_m
7	60.1	59.8
6.5	59.2	26.9 / 60.3
6	25.7 / 60.2	30.7 / 59.9
5	34.5 / 58.6	41.3 / 58.3

The melting profiles for both triplexes at pH 7 (Figure 6.5) showed monophasic curves almost identical to the one obtained for the target hairpin with slightly higher melting temperatures of 60.1 °C and 59.8 °C for the non-modified and modified triplexes, respectively. This suggests that at neutral pH, the DNA predominantly exists as a duplex and single strand, and little or no triplex is formed. This is most likely due to the lack of protonation of the third strand cytosine bases at neutral pH, which prevents the formation of Hoogsteen base pairs between the third strand and the purine strand of the duplex.
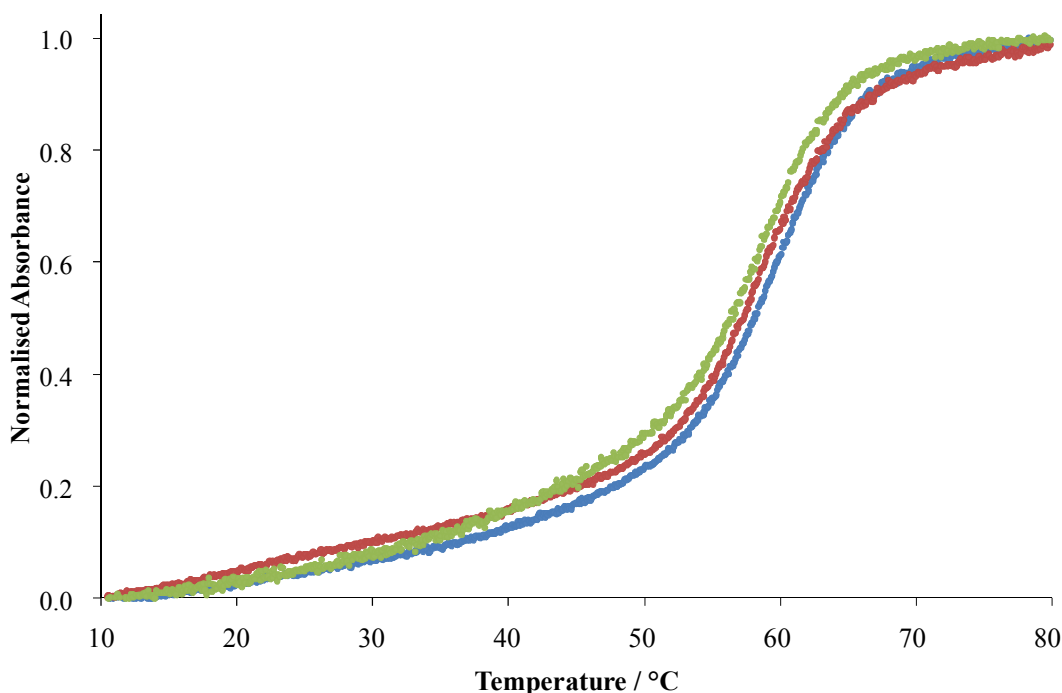
Figure 6.5. Normalised melting profiles for the non-modified triplex (blue) and modified triplex (red) at pH 7. The melting profile for the hairpin (green) is also shown for comparison. Absorbance was measured at 260 nm and samples were heated from 10 - 85 °C at a rate of 1 °C /min. Sample concentrations were 10 μM in 20 mM sodium phosphate buffer.

At pH 6, both triplexes produced clear biphasic melting curves (Figure 6.6) suggesting that pH 6 is sufficiently acidic to protonate the third strand cytosine iminos and the third strand is able to bind to the duplex by Hoogsteen base pairing. For both triplex structures a significant increase in T_m of approximately 9 °C (non-modified triplex) and 11 °C (modified triplex) was observed for the first melting transition (triplex to duplex unfolding) as the pH was decreased further from pH 6 to 5 (Figure 6.7), indicating that the triplex is more stable at a more acidic pH. This again is consistent with the requirement for protonation of cytosines in the third strand.
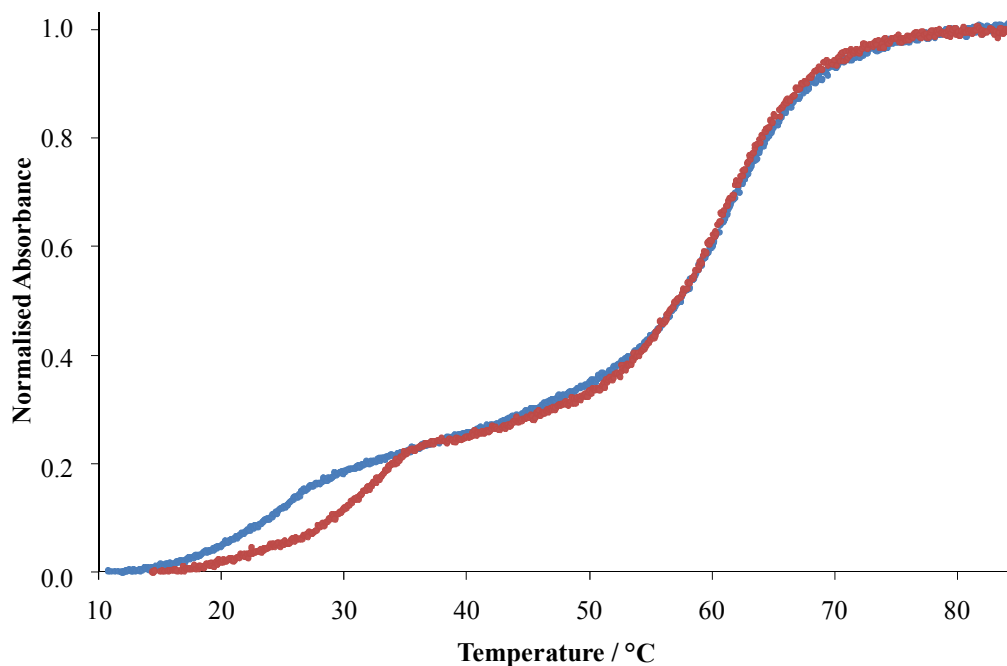
Figure 6.6. Normalised melting profiles for the non-modified triplex (blue) and modified triplex (red) at pH 6. Absorbance was measured at 260 nm and samples were heated from 10 - 85 °C at a rate of 1 °C /min. Sample concentrations were 10 μM in 20 mM sodium phosphate buffer.
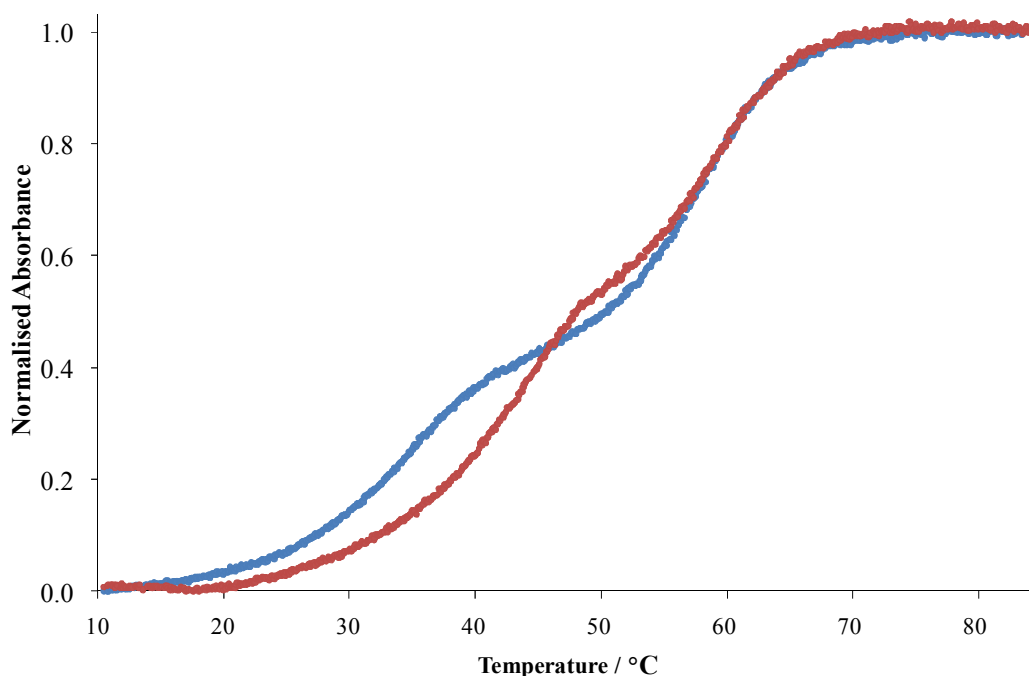
Figure 6.7. Normalised melting profiles for the non-modified triplex (blue) and modified triplex (red) at pH 5. Absorbance was measured at 260 nm and samples were heated from 10 - 85 °C at a rate of 1 °C /min. Sample concentrations were 10 μM in 20 mM sodium phosphate buffer.

From Figure 6.6, Figure 6.7 and Table 6.3 it can be seen that the first melting transition of the modified triplex, which corresponds to the triplex to duplex melting transition is shifted by 5 °C (pH 6) and 6.8 °C (pH 5) compared to the non-modified equivalent. This indicates that the triplex is significantly stabilised upon incorporation of two non adjacent 3′-S modifications within the third strand, particularly under more acidic conditions. In contrast, the second melting transition is unaffected; this is to be expected since it corresponds to unfolding of the hairpin, which contains no modifications and is unimolecular.

Since duplex formation is predominant at neutral pH, whereas at pH 6 triplex formation becomes favoured, additional melting curves were obtained for both triplex structures at a pH value of 6.5, in an attempt to narrow down the pH range for which the triplex is formed. The subsequent melting profiles are shown in Figure 6.8 and the melting temperatures given in Table 6.3.
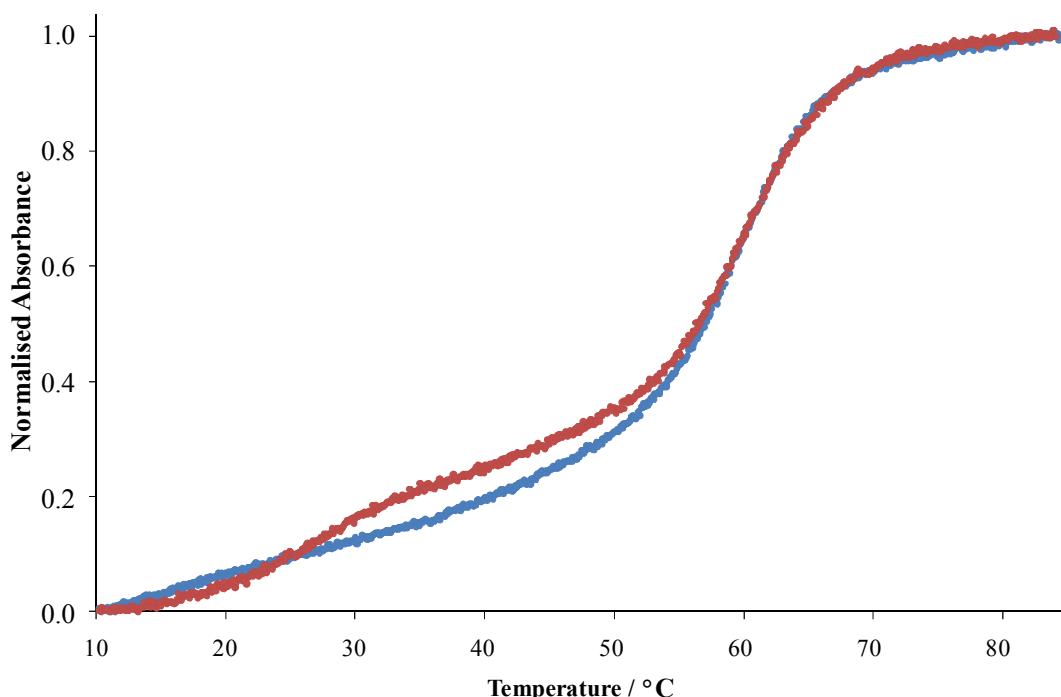
Figure 6.8. Normalised melting profiles for the non-modified triplex (blue) and modified triplex (red) at pH 6.5. Absorbance was measured at 260 nm and samples were heated from 10 - 85 °C at a rate of 1 °C /min. Sample concentrations were 10 μM in 20 mM sodium phosphate buffer.

At pH 6.5, the non-modified triplex exhibits a monophasic curve similar to the one obtained at pH 7, however the modified triplex melting profile is biphasic. This indicates that the modified triplex is prevalent at this pH whereas the non-modified triplex is not. Therefore the 3'-*S*-phosphorothiolate modifications increase the pH range for triplex formation and enables binding of the third strand to the hairpin at a pH value which is closer to physiological pH than for the non-modified strand.

Enthalpy contributions for the triplex melting curves cannot be calculated using the same method as used for the hairpin, since the triplex is not unimolecular and triplex unfolding is not a two state transition. Equation 6.5 was subsequently used to determine the van't Hoff enthalpies for each melting transition, which takes into

account the molecularity of the melting transition[172]. R denotes the universal gas constant, α represents the proportion of DNA in its folded state, and (∂ α/∂T) is given by the slope at the point around the T_m taken from a plot of fraction folded, α vs. T.

$$\Delta H_{vH} = RT_m{}^2 \left(\frac{\partial \propto}{\partial T} \right)$$
Equation 6.5

Each triplex melting profile was therefore converted into a fraction folded vs. temperature plot using Equation 6.1 and the slope around the melting temperatures was calculated (Figure 6.9).
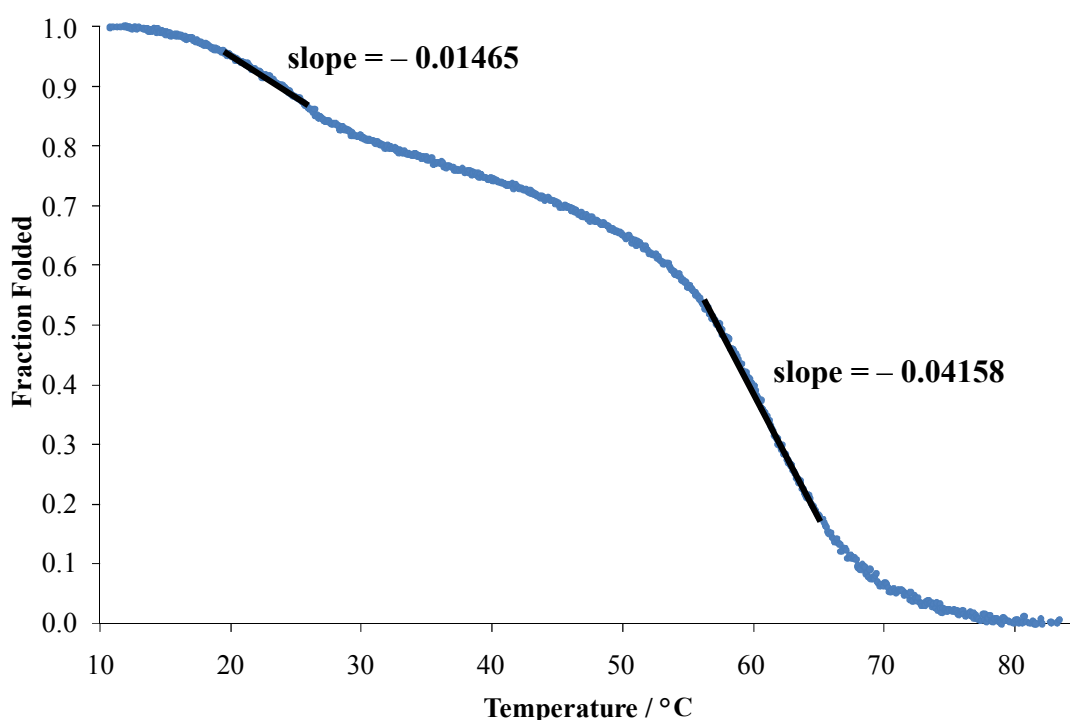
Figure 6.9. Fraction folded as a function of temperature for the non-modified triplex at pH 6.

Van't Hoff enthalpies were then determined for each unfolding transition for the non-modified and modified triplexes at each pH using Equation 6.5. Values are given in Table 6.4.

Table 6.4. Melting temperatures and van't Hoff enthalpies for non-modified and modified triplex formation calculated over a range of pH values. Melting temperatures are given in °C (± 0.5 °C), ΔH_{vH} are given in kJ mol^{-1} (± 10 %). (1) denotes triplex to duplex unfolding, (2) denotes duplex to single strand melting.

Triplex	pH	T_m (1)	ΔH_{vH} (1)	T_m (2)	ΔH_{vH} (2)
Non-modified	7	-	-	60.1	-203
	6.5	-	-	59.2	-180
	6	27.5	− 65	60.3	-154
	5	34.5	− 108	58.6	-138
Modified	7	-	-	59.8	-204
	6.5	26.9	− 57	60.3	-164
	6	30.7	− 88	59.2	-158
	5	41.3	− 151	58.8	-130

From Table 6.4 it can be seen that for each triplex, the duplex to single strand transition occurs with a more negative enthalpy than the subsequent removal of the third strand from the duplex. Assuming that ΔH_{vH} (1) and (2) are additive[162], the overall unfolding enthalpy for the non-modified triplex ranges from − 219 to − 246 kJ mol^{-1} between pH 6 and 5, which is more favourable than the unfolding enthalpy of − 86 kJ mol^{-1} measured for the hairpin. This confirms triplex formation within this pH range. Likewise, the overall unfolding enthalpy for the modified triplex ranges from − 221 to − 281 kJ mol^{-1} between pH 6.5 and 5, indicating that triplex formation is favoured over duplex formation within this pH range. The most favourable van't Hoff enthalpy of − 281 kJ mol^{-1} corresponds to the structure with the greatest stability, the modified triplex at pH 5.

6.6 Effect of salt on hairpin and triplex stability

100 mM sodium chloride was added to the hairpin sample and a UV melting profile was obtained in order to determine the effect of salt concentration on the stability of the structure. The melting curve is shown in Figure 6.10 and the melting temperatures are given in Table 6.5.
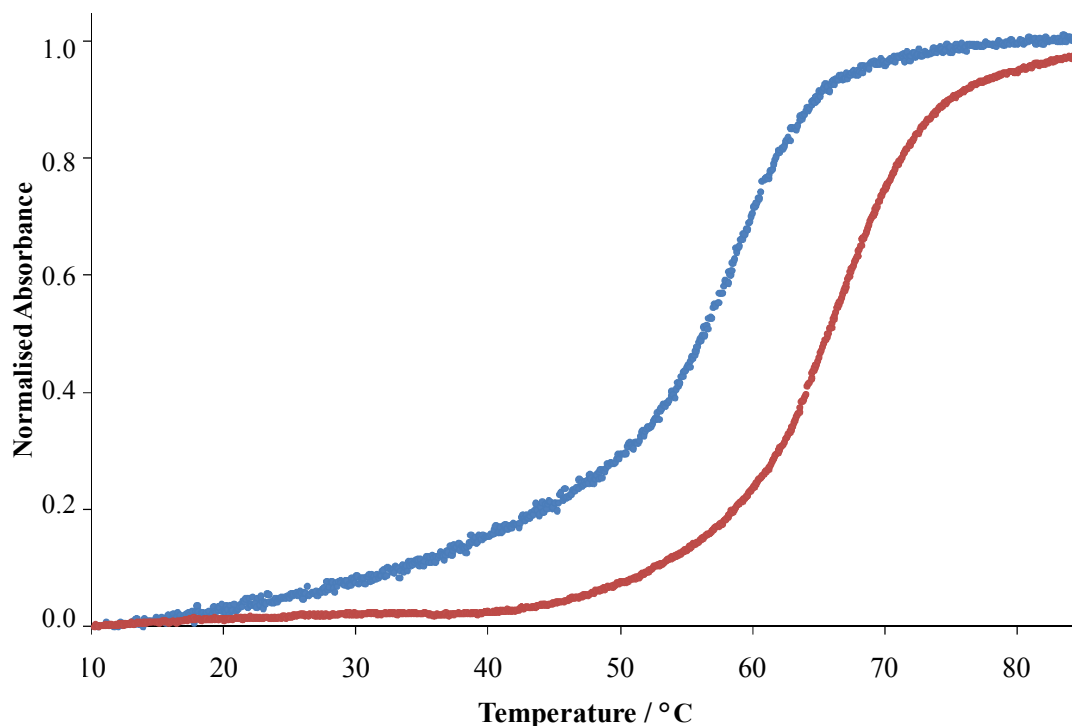
Figure 6.10. Normalised melting profiles for the hairpin at pH 7 in the presence and absence of 100 mM NaCl (red and blue curves, respectively). Absorbance was measured at 260 nm and samples were heated from 10 - 85 °C at a rate of 1 °C /min. Sample concentrations were 10 μM in 20 mM sodium phosphate buffer.

Table 6.5. Melting temperatures for the UV unfolding (260 nm) of the hairpin at pH 7 in the presence and absence of 100 mM NaCl. T_m values were averaged over two runs.

Salt concentration	T_m / °C
No salt	57.7
100 mM NaCl	66.2

The presence of NaCl stabilises the hairpin by 8.5 °C, which is consistent with previous reports on the effect of salt concentration on DNA duplex stability and can be attributed to minimisation of the electrostatic repulsions between the negatively charged phosphate groups of the strands[105, 173].

UV melting profiles were also obtained for the non-modified triplex at pH 7, 6 and 5 at the increased salt concentration of 100 mM NaCl and are shown in Figures 6.11 – 6.13, along with the melting temperatures which are given in Table 6.6.

Table 6.6. Melting temperatures for the UV unfolding (260 nm) of the non-modified triplex in the absence and presence of 100 mM NaCl. Values are given in °C.

pH	No salt	100 mM NaCl
7	60.1	68.0
6	25.7 / 60.2	21.2 / 66.6
5	34.5 / 58.6	30.6 / 64.1

The melting profiles for the non-modified triplex at pH 7 are very similar in appearance to those obtained for the hairpin, with a similar increase in T_m of 7.9 °C upon addition of salt. This provides further evidence that at neutral pH, the structure predominantly exists in the form of a hairpin and single strand.
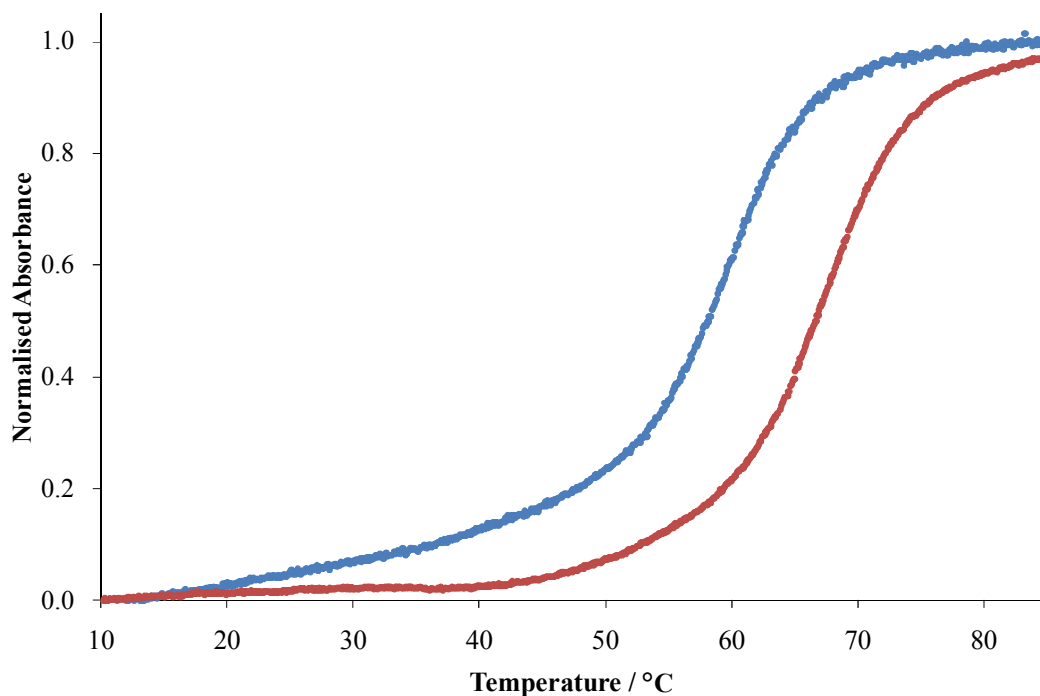
Figure 6.11. Normalised melting profiles for the non-modified triplex at pH 7 in the presence and absence of 100 mM NaCl (red and blue curves, respectively). Absorbance was measured at 260 nm and samples were heated from 10 - 85 °C at a rate of 1 °C /min. Sample concentrations were 10 μM in 20 mM sodium phosphate buffer.

At pH 6, biphasic melting curves are observed for the non-modified triplex both with and without added salt, which indicates that in both cases the triplex form is preferred over the duplex. The first melting transition experiences a decrease in T_m of − 4.5 °C at the increased salt concentration, whilst the melting temperature of the second transition is increased by 6.4 °C. Therefore whilst salt stabilises the duplex form, it has a destabilising effect on the triplex and causes the third strand to dissociate from the duplex at a lower T_m.
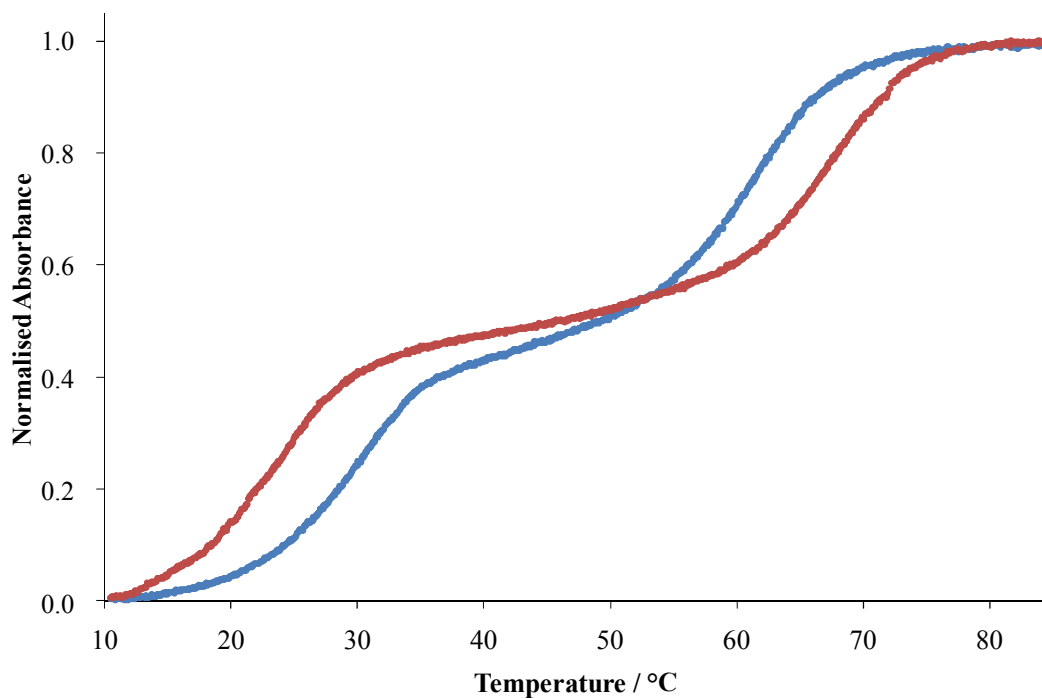
Figure 6.12. Normalised melting profiles for the non-modified triplex at pH 6 in the presence and absence of 100 mM NaCl (red and blue curves, respectively). Absorbance was measured at 260 nm and samples were heated from 10 - 85 °C at a rate of 1 °C /min. Sample concentrations were 10 μM in 20 mM sodium phosphate buffer.

A similar effect is also observed at pH 5, although the destabilisation of the triplex to duplex transition is not as pronounced compared to measurement made at pH 6. A slightly smaller decrease in T_m of – 3.9 °C is apparent. The second melting transition is stabilised by 5.5 °C upon the addition of 100 mM sodium chloride.
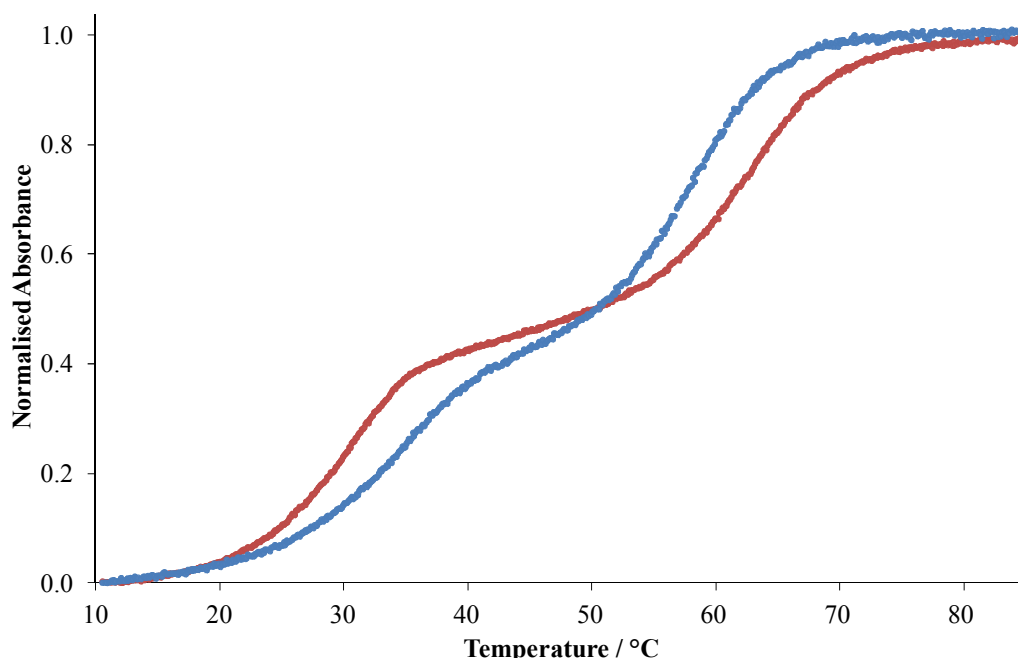
Figure 6.13. Normalised melting profiles for the non-modified triplex at pH 5 in the presence and absence of 100 mM NaCl (red and blue curves, respectively). Absorbance was measured at 260 nm and samples were heated from 10 - 85 °C at a rate of 1 °C /min. Sample concentrations were 10 μM in 20 mM sodium phosphate buffer.

The destabilisation of the triplex upon increasing the salt concentration is consistent with previous studies[159, 160]. This is because an uptake of protons is required for protonation of the cytosine bases in the third strand, and as the Na^+ concentration increases, more protons are required to push the equilibrium towards the favourable formation of a triplex[160, 162, 174]. This results in a decrease in the pK_a of the triplex and cytosine protonation is only induced at a lower pH. This effect outweighs the charge stabilisation and minimisation of electrostatic repulsions between the negatively charged phosphate backbones of the three strands.

6.7 Discussion

UV thermal analysis of the hairpin was carried out and melting of the structure was found to be independent of pH over the range 5 - 7. The melting temperature was obtained by two different methods, firstly by measuring the maximum value of the first derivative curve and secondly by extrapolating the fraction folded vs. temperature plot at the point where exactly half of the hairpin molecules were in their folded state. A slightly higher T_m of 57.7 °C was obtained using the derivative method compared to 56.7 °C which was measured from the fraction folded vs. temperature plot. The latter method for calculating T_m is known to give more accurate values as the baseline slope can affect the maximum of the derivative curve, resulting in over- or underestimation of the T_m value[171].

Thermodynamic parameters were also calculated for folding of the hairpin and the relative size of the parameters indicates that the process is enthalpy driven, with a small entropy contribution. At 37 °C, negative $\Delta G°$ values were obtained which confirms that hairpin folding is favourable at physiological temperature.

Melting of the non-modified triplex over the pH range 5 - 7 was also monitored by UV. Shape analysis of the subsequent melting curves indicated that the triplex does not form under physiological conditions (i.e. pH 7), which is consistent with literature precedence regarding the requirement for cytosine protonation in the third strand. At acidic pH, biphasic curves were produced which confirms triplex formation, and also indicates that the structure melts in two separate transitions. The first melting transition which corresponds to removal of the third strand from the hairpin was found to be dependent on pH, with the highest T_m value recorded at pH 5. This indicates that the triplex is more stable at more acidic pH which again is

in agreement with the literature. The second melting transition which corresponds to unfolding of the hairpin into a random coil was pH independent as expected for melting of an intramolecular system.

UV thermal melting profiles of the $3'$-S-phosphorothiolate modified triplex as a function of pH were also obtained. As with the non-modified equivalent, triplex formation did not occur at pH 7 however the modification resulted in a stabilisation of the triplex structure at acidic pH as indicated by an increased T_m value by up to 6.8 °C for the first transition of the biphasic melting profiles. This stabilisation was a result of altering the sugar pucker of the deoxyribose rings attached to, and $3'$ to the modification as determined in Chapters 4 and 5.

In addition, the upper pH limit for triplex formation was increased upon incorporation of the modifications into the third strand, as indicated by the monophasic and biphasic appearance of the melting profiles for the non-modified and modified triplexes, respectively obtained at pH 6.5. This could have profound implications for the potential use of triplex forming oligonucleotides in antigene therapies, which would require formation of a stable triplex at physiological pH.

Van't Hoff enthalpies were also calculated for the unfolding of both triplexes, and both triplexes were found to unfold with larger enthalpies than the hairpin, which is consistent with the incorporation of a third strand[160]. In addition, unfolding of the modified triplex proceeds with a larger enthalpy compared with that of the non-modified triplex, confirming that the inclusion of the $3'$-S-phosphorothiolate linkages results in stabilisation of the structure. Whilst the ΔH_{vH} enthalpies calculated from the UV melting curves give a good indication of the unfolding thermodynamics, erroneous values can be obtained due to additional intermediate states and

equilibria[171]. Therefore these should be interpreted with caution and so large errors of ± 10 % were applied. More accurate thermodynamic parameters may be obtained from differential scanning calorimetry experiments; however this is beyond the scope of the project.

Additional melting profiles were obtained for the hairpin and non-modified triplex at an increased salt concentration of 100 mM NaCl. The presence of salt was found to have a stabilising effect on the hairpin, with an increase in T_m of 8.5 °C being observed. This is consistent with previous reports and can be attributed to stabilisation of the negatively charged phosphate groups by Na^+ ions. In contrast, salt was found to have a destabilising effect on the triplex structure. Whilst Na^+ would minimise charge repulsion between the three phosphate backbones of the three strands of the triplex, this effect is overridden by an exclusion of counterions due to cytosine protonation. A decrease in T_m of − 4.5 °C upon the addition of salt was observed at pH 6, with a smaller destabilisation of − 3.9 °C measured at pH 5. It is likely that at the lower pH, a decreased release of counterions is observed and so the salt concentration has a lesser effect on the stability of the structure[160, 162].

6.8 Conclusions

UV thermal melting analysis of the target hairpin and both the non-modified and modified triplexes was performed. Unfolding of the hairpin was found to be independent of pH, whereas triplex unfolding was determined to be pH dependent, with the most stable triplexes observed at acidic pH.

The incorporation of two 3′-*S*-phosphorothiolate linkers into the third strand of the triplex was found to have a stabilising effect on the structure at acidic pH. The

modifications also increased the upper pH limit for triplex formation from pH 6 to pH 6.5.

The effect of salt concentration on the molecules was also analysed, with an increased NaCl concentration resulting in destabilisation of the triplex and stabilisation of the duplex.

Chapter 7

Conclusions

Chapter 7. Conclusions

Triple helical nucleic acids have been implicated in a diverse range of biological processes and functions; one particularly promising application being the use of triple helical DNA in regulating the process of transcription, the so-called antigene technology. Homopurine-homopyrimidine tracts of DNA, which are required for triplex formation, are overexpressed in genes associated with several diseases such as HIV[80], diabetes[60] and cancer[72]. It is therefore of interest to design TFOs which form highly stable triplexes when bound to a DNA duplex; these could potentially prevent DNA replication and, ultimately, gene expression when utilised as therapeutic agents. This project was aimed at probing the structure and stability of a specific naturally occurring DNA triplex and a subtly modified equivalent.

The modification chosen was a 3′-S-phosphorothiolate link, in which the 3′ oxygen atom on a deoxyribose sugar is replaced with a less electronegative sulfur atom. Two non-adjacent 3′-S-phosphorothiolate linkages were incorporated into the third strand of the triplex in question, using NMR spectroscopy and UV thermal melting analysis techniques to determine the consequences of the modification. It was anticipated that the modification would result in a change in conformation of the sugar ring to which it was attached and, to a lesser extent, the sugar 3′ to it[92, 97, 168], from a DNA-like south conformation to a north pucker which is more typical of RNA nucleotides. It is known that triplexes comprising a DNA duplex and an RNA third strand exhibit enhanced stability compared to their all DNA counterparts and so it was hoped that the modification, which acts as an RNA mimic, would result in stabilisation of the triplex.

The modified triplex was studied alongside the natural DNA equivalent and the target hairpin was analysed prior to binding of the TFO in order to fully characterise the effects of the 3′-S-phosphorothiolate on the triplex formation, structure, and stability.

7.1 Target hairpin

It was necessary to elucidate the structure and thermal properties of the target hairpin prior to any studies of the triplex in order to characterise structural changes to the hairpin induced by binding of the third strand. A biologically relevant homopurine-homopyrimidine sequence connected by a triethylene glycol linker was chosen, synthesised and analysed by NMR spectroscopy and UV spectrophotometry. The structure of the hairpin was observed to be a typical B-type DNA-like helical structure with Watson-Crick base pairing between the two strands. As expected for a B-type helix, sugar puckers were all found to reside in the south conformation and *anti*-glycosidic bond orientation. High resolution, low energy structures were then generated based on the NMR data obtained and structural features were all consistent with a right handed, B-type DNA helix. The hairpin was found to have a melting temperature (defined as the temperature at which exactly half of the sample is folded and the other half is dissociated into a random coil) of approximately 57 °C, and melting was found to be independent of pH over the range 5 - 7.

7.2 Native triplex

A complementary TFO composed of all natural bases was synthesised and annealed with the hairpin in a 1:1 ratio. As anticipated, the third strand bound to the purine

strand of the hairpin in a parallel orientation by Hoogsteen hydrogen bonds. Binding of the natural TFO appeared to have little effect on the structure of the target hairpin, which implied that the third strand is readily accommodated in the major groove, causing little perturbation to the structure in doing so. All sugar puckers analysed for both the hairpin and third strand residues were found to reside in the south conformation, signifying an overall B-type structure. The triplex was only stable at acidic pH and melted in two distinct transitions indicating that removal of the third strand and unfolding of the hairpin do not occur simultaneously.

7.3 3′-*S*-phosphorothiolate modified triplex

Two non adjacent 3′-*S* linkers were incorporated in the third strand of the triplex and the effects of the modification could then be studied, with reference to the natural equivalent. As anticipated and in agreement with previous studies[168], the modifications caused a conformational shift in both the modified and 3′ sugar rings from a DNA-like to an RNA-like conformation. These effects were localised, with the remaining sugar puckers along the length of the sequence remaining unchanged. Changes were noted in the chemical shifts of the protons within the modified triplex compared to those within the native triplex; however no significant changes were observed in the structural analysis, implying that any global effects of the modifications are subtle. UV thermal melting studies showed that the localised conformational changes resulted in an increase in binding affinity of the TFO to the hairpin, with an observed increase in melting temperature of up to 6.8 °C (depending on pH). The modifications also increased the upper pH limit for triplex formation by

half a unit to pH 6.5, however, even with the modifications in the third strand, the triplex remained unstable at physiological pH.

7.4 Future work

The aim of this project - to analyse the effects of incorporating two 3′-*S*-phosphorothiolate modifications in to the third strand of a DNA triplex - has been accomplished. However, there is plenty of scope for further research.

It would be worth following up the thermodynamic data with additional experiments. Isothermal titration calorimetry, in which the third strand would be titrated into the hairpin solution could give information regarding the concentration dependence and binding affinity of the third strand for the hairpin. Differential scanning calorimetry experiments could enable more accurate thermodynamic parameters to be calculated and provide an insight into the unfolding mechanism.

It would be useful to examine an equivalent triplex which comprises a DNA target hairpin with an unmodified RNA third strand. This would be useful for comparison to the modified DNA triplex, and to determine to what extent the 3′-*S* modifications mimic RNA.

The homopurine-homopyrimidine requirement for triplex formation severely limits their potential use, and so future work may also involve employing 3′-*S* nucleotides in conjunction with a different modification aimed at reducing the need for long stretches of polypurine-polypyrimidine sequences.

If triplex forming oligonucleotides are to be used as therapeutic agents they must form stable triplexes at physiological pH, therefore the acidic pH requirement

associated with parallel triplexes must be overcome. Incorporating 3′-*S*-phosphorothiolate modifications into the TFO reduced the upper pH dependence for triplex formation by half a unit, which is a step in the right direction. However, neither the natural nor modified triplex was stable at physiological pH and so further stabilisation is needed before this modification can be considered for use within oligonucleotide-based drugs. As the modifications were positioned only three residues apart, it would be useful to examine the effects of different positioning of the 3′-*S* linkages along the length of the oligonucleotide chain to see if any further stabilisation can be achieved. Likewise, increasing the number of modifications would be of interest and feasible, since synthesis of the 3′-*S*-phosphorothiolate nucleotides is relatively straightforward and coupling efficiencies have improved considerably in recent years. Additional biological information is required in order for the therapeutic potential of the 3′-*S* modification within triplex forming oligonucleotides to be fully realised, all of which is beyond the scope and expertise of this laboratory. This project has, however, provided a basis for investigating the effects of 3′-*S*-phosphorothiolate linkers within triplex forming oligonucleotides, and the promising results achieved present a good starting point for further research.

Chapter 8

Experimental

Chapter 8. Experimental

8.1 Oligonucleotides

8.1.1 DNA synthesis

Oligonucleotide synthesis followed the standard phosphoramidite route[175]. DNA strands were synthesised on MerMade 4 solid phase oligonucleotide synthesiser using the standard DMT-on automated synthesis protocol. Standard A, C, G and T phosphoramidites as well as the dU phosphoramidite, spacer-9 linker, controlled pore glass (CPG) columns and other standard synthesis reagents were purchased from Link Technologies Ltd. (Lanarkshire, Scotland). Phosphoramidites were dissolved in anhydrous acetonitrile under argon, to give a concentration of 0.1 M. Cleavage of the DNA strands from the CPG was then achieved by suspending at 55 °C for 18 hours in fresh ammonium hydroxide solution, which was prepared by bubbling ammonia gas through RNAse free water for 1 hour. The ammonium hydroxide solution was then removed by blowing down to minimum solvent with nitrogen gas and finally, the samples were dissolved in 1 ml RNAse free water. Samples were then stored at 4 °C before purification.

The 3'-S-phosphorothiolate modified third strand was synthesised by Dr. Inder Bhamra and Alannah Clelland at the University of Liverpool, using a method described in the literature[91].

8.1.2 Oligonucleotide purification

The oligonucleotides were purified by reverse phase High Performance Liquid Chromatography (HPLC). The crude sample was filtered using a 0.45 µm syringe filter to remove any remaining CPG beads and then diluted by a factor of 50. For the hairpin and non-modified third strand, HPLC was performed using a Dionex P580 pump with a Gynotek Degasys DG-2410 system and a Gynotek UVD 340S diode array UV detector. The system was controlled by a PC using Chromeleon 6.8 software[176]. Analytical and preparatory runs were carried out on a 4 x 250 mm RP18 column with a 5 µm particle size (Merck LiChroCart 250-4) heated to 55 °C. Two buffer solutions were used; buffer A was 100 mM ammonium acetate and buffer B was acetonitrile. The sample was loaded onto the column via a 500 µl injection loop and the elution of the DNA was followed by monitoring the UV absorbance at 260 nm. Desired fractions were collected manually. During each run the composition of the eluent was gradually changed from 100 % buffer A (0 % buffer B) to 90 % buffer A (10 % buffer B) over 25 minutes, then to 55 % buffer A (45 % buffer B) over the next 5 minutes, and finally back to 100 % buffer A (Figure 8.1). Each preparatory run lasted 45 minutes.
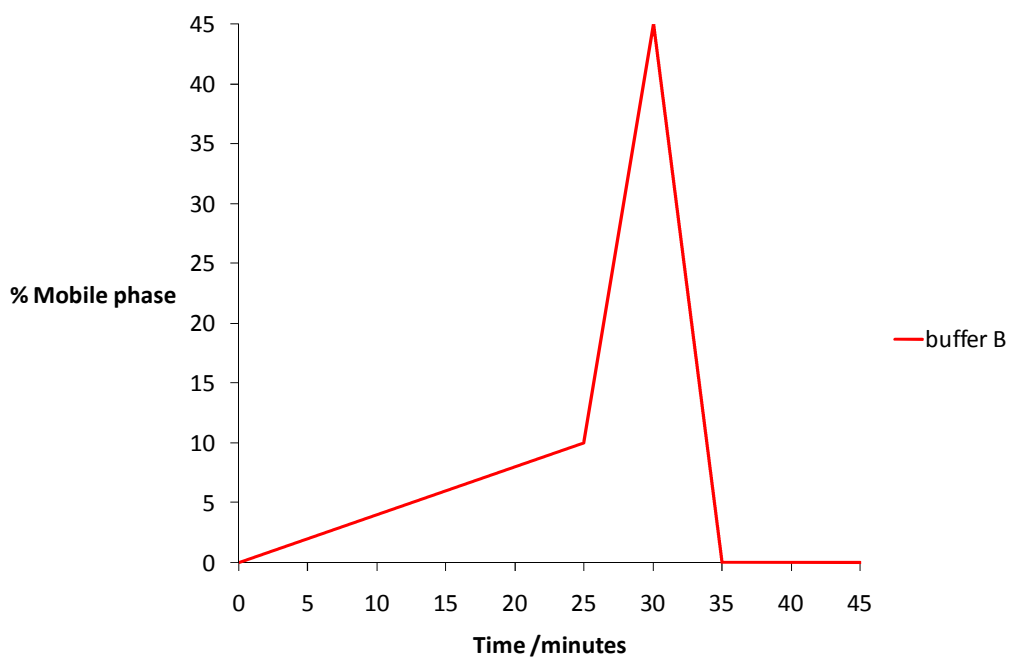
Figure 8.1. Graph illustrating the eluent gradient system used for reverse phase HPLC of the oligonucleotides.

For the modified third strand, an alternative solvent system was used using triethylamine bicarbonate (TEAB) buffer solutions. Buffer A was 0.1 M TEAB solution and buffer B was 0.1 M TEAB solution containing 40 % acetonitrile. During each run the composition of the eluent was gradually changed from 100 % to 0 % buffer A (0 % to 100 % buffer B) over 20 minutes and held for a further 5 minutes, then back to 100 % buffer A over a subsequent 20 minute period (Figure 8.2).
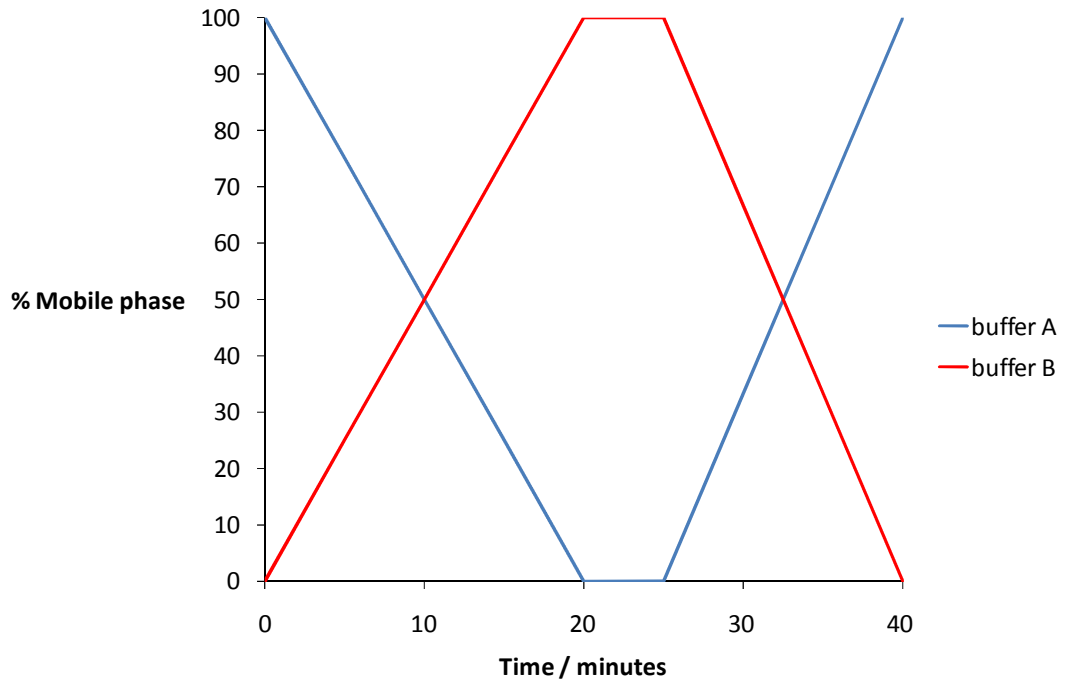
Figure 8.2. Graph illustrating the alternative eluent gradient system used for reverse phase HPLC of the oligonucleotides.

8.1.3 Removal of the DMT protecting group

The DMT group was removed from the DNA strands by adding aqueous acetic acid solution (80 %, 0.5 ml) to the samples, shaking and leaving to stand for one hour. The acid was then removed *in vacuo* and the strands were dissolved in RNAse free water (0.5 ml). 3 x 0.5 ml extractions with ethyl acetate were carried out and the aqueous layers retained, combined and concentrated *in vacuo*. The fully deprotected oligonucleotides were then re-suspended in RNAse free water.

8.1.4 Desalting procedure

NAP25 columns were purchased from GE healthcare (Chalfont St. Giles, UK) were washed with 25 ml RNAse free water before adding the DNA samples (2 ml per

column). The columns were eluted with RNAse free water to remove any salt present and 10 x 1 ml fractions were collected in sterilised eppendorfs. The UV absorbance at 260 nm of each fraction was measured using a Hewlett Packard 8452A diode array UV/Visible spectrometer. The fractions which exhibited a significant absorbance were retained. 5 μl sodium phosphate buffer solution was then added to each desalted DNA fraction before snap freezing in liquid nitrogen and placed under vacuum overnight until completely dry. The fractions were then combined by re-dissolving in D_2O.

8.1.5 Calculation of oligonucleotide concentration

The concentration of each oligonucleotide sample was calculated by measuring the UV absorbance at a wavelength of 260 nm and using the Beer-Lambert law (Equation 8.1).

$$A_{260} = \varepsilon_{260} c l$$ **Equation 8.1**

A_{260} denotes the UV absorbance at 260 nm, ε represents the extinction coefficient which was calculated based on a method in the literature[105], c denotes the concentration of the sample and l is the path length through which the light travels, which was given by the diameter of the cuvette (10 mm). 10 μl of each DNA solution was made up to 1 ml in RNAse free water, and the absorbance was measured using a Hewlett Packard 8452A diode array UV/Visible spectrophotometer. The extinction coefficients were calculated as: ε (hairpin) = 235970 M^{-1} cm^{-1}, ε (single strands) =

94000 M^{-1} cm^{-1}. The volumes needed for a 1:1 ratio of hairpin to single strand which was required for the triplex samples was subsequently calculated from the strand concentrations. The two strands were then combined and freeze dried.

8.1.6 NMR sample preparation

The DNA samples were freeze dried and dissolved in 1 ml D_2O that was purchased from Sigma-Aldrich (Dorset, England) an additional three times to allow hydrogen-deuterium exchange. 600 µl of 20 mM phosphate buffer (pH 7) was prepared in D_2O (see section 8.1.7) then freeze dried and combined with the DNA sample in 99.9% ampule D_2O (600 µl), or 9:1 H_2O:D_2O (540 µl RNAse free water, 60 µl ampule D_2O) depending on the experiment requirements. The pH of the triplex samples was adjusted to pH 5 by adding a small amount of deuterium chloride, purchased from Goss Scientific (Cheshire, England). Samples were then annealed by heating to 95 °C for 5 minutes and left to cool slowly to room temperature, to allow sufficient time for the lowest energy conformation to form. The sample was then transferred to a clean NMR tube using a sterilised pipette.

8.1.7 Preparation of sodium phosphate buffer

Na_2HPO_4 and NaH_2PO_4 were purchased from Sigma Aldrich (Dorset, England). 0.2 M solutions were then prepared by dissolving 0.28 g Na_2HPO_4 and 0.24 g NaH_2PO_4 each in 10 ml D_2O. 0.2 M phosphate buffer (pH 7) was then prepared by combining 4.9 ml 0.2 M Na_2HPO_4 and 5.1 ml 0.2 M NaH_2PO_4. This was then diluted by a factor of 10 to give 20 mM phosphate buffer solution.

8.2 NMR experiments

NMR experiments were performed using a Varian Unity Inova Spectrometer with a ^{1}H operating frequency of either 500 or 750 MHz and a 5 mm [1]H {15N-13C} probe. Standard 5 mm Norell S500 NMR tubes containing 600 µl of sample were used. Experiments were run at 20 °C unless otherwise stated.

8.2.1 Varian 500 MHz ^{1}H NMR

1 presaturation experiments were recorded for the hairpin, non-modified triplex and modified triplex with a saturating pulse applied for 1.5 seconds at a power of 2 dB. A spectral width of 6000 Hz was collected in 8192 pairs of data points and experiments were recorded using 256 scans and a recycle delay of 1 second. A high power pulse width of 3 µs (equivalent to a 30 ° angle) was employed for all presat experiments.

DQF-COSY, TOCSY and NOESY experiments were recorded for the hairpin, non-modified and modified triplex samples in D_2O. A spectral width of 6000 Hz in both dimensions was collected in 4096 pairs of data points in t_2. The 90x ° pulse width differed for each sample but was generally around 8 µs at a power of 60 dB. The DQF-COSY experiments used 64 scans and 256 t_1 increments and for the TOCSY and NOESY spectra 128 scans and 256 t_1 increments were used. For the TOCSY a spinlock time of 80 ms was employed and an MLEV-17 pulse train was applied using a power of 45 dB. For the NOESY experiments used in proton resonance assignments mixing times of 200 and 400 ms was used; additional NOESY spectra with mixing times of 50, 100, 150, 250, 300 and 350 ms were obtained for use in structure generation of the hairpin. A 1 second presaturation pulse was applied for all

experiments in order to saturate the water signal, which was also suppressed during the mixing time.

8.2.2 Varian 750 MHz ^{1}H NMR

NOESY spectra were recorded at 750 MHz for the hairpin, non-modified triplex and modified triplex in D_2O to aid with proton resonance assignments. A spectral width of 6000 Hz was collected in 4096 pairs of data points at a power of 56 dB. 64 scans, 256 increments and a mixing time of 200 ms were used. Additional NOESY spectra were acquired for the hairpin sample at temperatures of 25 and 30 °C to assist with assignments. Spectra were also recorded at 750 MHz for each oligonucleotide in a 9:1 H_2O: D_2O solvent mix using a NOESY_dpfgse pulse sequence at 2 °C. Spectral widths of 15000 Hz and 18000 Hz for the hairpin and triplex samples respectively, were collected in 4096 pairs of data points at a power of 56 dB. 128 scans, 256 increments and a mixing time of 200 ms were used.

8.2.3 Data processing

All one dimensional data were processed using ACD/SpecManager 12.0 software[177]. All two dimensional datasets were processed using VNMR 6.1C software (Varian, Inc.) and spectra were assigned using SPARKY 3.111[178]. All spectra were referenced to the residual water peak (4.8 ppm).

8.3 Structure calculations

8.3.1 Molecular modelling

Molecular modelling was performed using the CNS_solve package, version 1.1 on a UNIX platform. Simulated annealing using torsion angle dynamics was employed, consisting of one high temperature heating stage followed by two cooling stages and 1000 steps of conjugate gradient energy minimisation. Minimisation and conformational search was carried out using the AMBER[179] force field which is a well parameterised force field for biopolymers. Implicit waters were included as the solvent so that the energy minimisation was comparable to experimental conditions. Constant dielectric ($\varepsilon = 80$) treatment was used to account for effects of water on the electrostatic interactions. Sodium counterions were also included in each model. Simulated annealing was initially performed using unfolded starting structures; subsequent rounds were then executed using the lowest energy structure generated from the previous round as a starting point. Rounds of annealing were repeated until acceptable structures were consistently being generated.

Topology and parameter files were modified for the triplex structures to include atom masses, charges, bond lengths and bond angles for the protonated cytosine residues in the third strand. Charges for the atoms on the C^+ residues were provided by Charles Laughton at the University of Nottingham. 3′-*S*-phosphorothiolate thymidine residues were also created for the modified triplex structures. Sulfur atom masses and charges, P-S-C bond lengths, and bond angles were all included in the DNA topology and parameter files.

8.3.2 Analysis of structures

Structures were visualised in PyMOL[180] and distances between atoms were measured using the wizard function. Backbone and helical parameter analysis was carried out by analysing the pdb file in CURVES^{+141} web server and major groove widths were calculated using the 3DNA[148] web server package. For the triplex structures, the hairpin and third strands were examined separately as three stranded nucleic acid structures could not be analysed in CURVES^{+} or 3DNA.

8.4 UV thermal melting studies

8.4.1 Sample preparation

Oligonucleotide samples were diluted to approximately 1.2 ml in 20 mM sodium phosphate buffer, and the pH of each sample was adjusted to the required pH using deuterium chloride and sodium deuteroxide. Samples were annealed prior to UV analysis.

8.4.2 UV melting experiments

UV melting profiles for all DNA structures were obtained at the University of Bradford using a Varian: Cary 400 Bio UV Vis Spectrophotometer version 8.01 and Thermal software version 3.0. A reference sample of 20 mM buffer solution was heated alongside the DNA samples and the absorbance at 260 nm was monitored. Samples were heated between 10 °C and 85 °C at a rate of 1 °C per minute, with readings being taken at 0.1 °C intervals. Two runs were performed for each sample, with an additional third run if there was a difference in T_m of > 0.2 ° C.

Appendix

Appendix

A1. NOe Distance restraints files for target hairpin

200 ms non-exchangeable proton distance restraints file

assign (resid 1 and name H2' and segid 1) (resid 1 and name H1' and segid 1) 2.2 1.1 1.1
assign (resid 1 and name H2' and segid 1) (resid 1 and name H8 and segid 1) 2.9 1.8 1.8
assign (resid 1 and name H2" and segid 1) (resid 1 and name H4' and segid 1) 2.5 1.5 1.5
assign (resid 1 and name H2" and segid 1) (resid 1 and name H1' and segid 1) 2.5 1.3 1.3
assign (resid 1 and name H1' and segid 1) (resid 1 and name H8 and segid 1) 3.1 1.6 1.6
assign (resid 1 and name H2' and segid 1) (resid 1 and name H2" and segid 1) 2.0 1.0 1.0
assign (resid 1 and name H2" and segid 1) (resid 1 and name H8 and segid 1) 3.2 1.6 1.6
assign (resid 1 and name H4' and segid 1) (resid 1 and name H1' and segid 1) 3.7 1.9 1.9
assign (resid 1 and name H4' and segid 1) (resid 1 and name H8 and segid 1) 3.2 1.6 1.6
assign (resid 1 and name H5' and segid 1) (resid 1 and name H8 and segid 1) 3.6 1.8 1.8
assign (resid 1 and name H5" and segid 1) (resid 1 and name H4' and segid 1) 1.7 0.9 0.9
assign (resid 2 and name H8 and segid 1) (resid 1 and name H8 and segid 1) 4.4 2.2 2.2
assign (resid 2 and name H2' and segid 1) (resid 1 and name H8 and segid 1) 3.3 1.6 1.6
assign (resid 2 and name H1' and segid 1) (resid 1 and name H8 and segid 1) 3.1 2.3 2.3
assign (resid 2 and name H1' and segid 1) (resid 2 and name H8 and segid 1) 3.0 1.5 1.5
assign (resid 2 and name H2' and segid 1) (resid 2 and name H1' and segid 1) 2.4 1.0 1.0
assign (resid 2 and name H2' and segid 1) (resid 2 and name H2" and segid 1) 2.2 0.9 0.9
assign (resid 2 and name H2' and segid 1) (resid 2 and name H8 and segid 1) 2.4 2.0 2.0
assign (resid 2 and name H2" and segid 1) (resid 2 and name H1' and segid 1) 2.7 1.1 1.1
assign (resid 2 and name H4' and segid 1) (resid 2 and name H1' and segid 1) 3.1 1.2 1.2
assign (resid 2 and name H5' and segid 1) (resid 2 and name H4' and segid 1) 1.9 0.8 0.8
assign (resid 2 and name H5' and segid 1) (resid 2 and name H8 and segid 1) 3.3 3.0 3.0
assign (resid 2 and name H2' and segid 1) (resid 2 and name H4' and segid 1) 4.6 1.8 1.8
assign (resid 2 and name H2" and segid 1) (resid 2 and name H8 and segid 1) 3.2 1.3 1.3
assign (resid 2 and name H4' and segid 1) (resid 2 and name H8 and segid 1) 4.8 1.9 1.9
assign (resid 3 and name H1' and segid 1) (resid 2 and name H8 and segid 1) 3.8 2.0 2.0
assign (resid 3 and name H2' and segid 1) (resid 2 and name H8 and segid 1) 2.9 2.0 2.0
assign (resid 2 and name H8 and segid 1) (resid 3 and name H8 and segid 1) 3.9 2.5 2.5
assign (resid 3 and name H2' and segid 1) (resid 3 and name H1' and segid 1) 2.9 1.2 1.2
assign (resid 3 and name H1' and segid 1) (resid 3 and name H8 and segid 1) 3.1 1.2 1.2
assign (resid 3 and name H2' and segid 1) (resid 3 and name H2" and segid 1) 2.5 1.0 1.0
assign (resid 3 and name H2' and segid 1) (resid 3 and name H8 and segid 1) 2.4 1.5 1.5
assign (resid 3 and name H2" and segid 1) (resid 3 and name H1' and segid 1) 2.9 1.2 1.2
assign (resid 3 and name H2' and segid 1) (resid 3 and name H4' and segid 1) 3.4 1.4 1.4
assign (resid 3 and name H2" and segid 1) (resid 3 and name H8 and segid 1) 3.5 1.4 1.4
assign (resid 3 and name H3' and segid 1) (resid 3 and name H1' and segid 1) 3.7 1.5 1.5
assign (resid 3 and name H4' and segid 1) (resid 3 and name H8 and segid 1) 3.3 2.5 2.5
assign (resid 3 and name H4' and segid 1) (resid 3 and name H1' and segid 1) 2.7 1.1 1.1
assign (resid 3 and name H8 and segid 1) (resid 4 and name H8 and segid 1) 4.0 3.0 3.0
assign (resid 4 and name H2' and segid 1) (resid 4 and name H2" and segid 1) 2.2 0.9 0.9
assign (resid 4 and name H2' and segid 1) (resid 4 and name H3' and segid 1) 2.5 1.0 1.0
assign (resid 4 and name H2' and segid 1) (resid 4 and name H4' and segid 1) 3.3 1.3 1.3
assign (resid 4 and name H2' and segid 1) (resid 4 and name H8 and segid 1) 2.6 1.5 1.5
assign (resid 4 and name H2" and segid 1) (resid 4 and name H1' and segid 1) 2.3 0.9 0.9
assign (resid 4 and name H2" and segid 1) (resid 4 and name H3' and segid 1) 2.2 0.9 0.9
assign (resid 4 and name H2" and segid 1) (resid 4 and name H4' and segid 1) 2.7 1.5 1.5
assign (resid 4 and name H5' and segid 1) (resid 4 and name H4' and segid 1) 2.0 0.8 0.8

assign (resid 4 and name H1' and segid 1) (resid 4 and name H8 and segid 1) 3.8 1.5 1.5
assign (resid 4 and name H3' and segid 1) (resid 4 and name H1' and segid 1) 4.0 1.6 1.6
assign (resid 4 and name H5" and segid 1) (resid 4 and name H4' and segid 1) 2.0 0.8 0.8
assign (resid 4 and name H4' and segid 1) (resid 4 and name H8 and segid 1) 3.7 2.0 2.0
assign (resid 4 and name H2' and segid 1) (resid 4 and name H1' and segid 1) 2.5 1.0 1.0
assign (resid 4 and name H2" and segid 1) (resid 4 and name H8 and segid 1) 2.8 1.5 1.5
assign (resid 4 and name H4' and segid 1) (resid 4 and name H1' and segid 1) 3.1 1.2 1.2
assign (resid 4 and name H5" and segid 1) (resid 4 and name H8 and segid 1) 3.7 3.0 3.0
assign (resid 4 and name H5' and segid 1) (resid 4 and name H8 and segid 1) 3.9 2.5 2.5
assign (resid 4 and name H3' and segid 1) (resid 4 and name H8 and segid 1) 3.5 2.0 2.0
assign (resid 5 and name H8 and segid 1) (resid 4 and name H8 and segid 1) 4.5 1.8 1.8
assign (resid 5 and name H1' and segid 1) (resid 4 and name H8 and segid 1) 3.4 2.0 2.0
assign (resid 5 and name H2' and segid 1) (resid 4 and name H8 and segid 1) 3.2 1.3 1.3
assign (resid 5 and name H2" and segid 1) (resid 4 and name H8 and segid 1) 3.2 2.0 2.0
assign (resid 5 and name H2' and segid 1) (resid 5 and name H1' and segid 1) 2.9 1.2 1.2
assign (resid 5 and name H2" and segid 1) (resid 5 and name H1' and segid 1) 2.6 1.0 1.0
assign (resid 5 and name H4' and segid 1) (resid 5 and name H1' and segid 1) 3.1 1.2 1.2
assign (resid 5 and name H5" and segid 1) (resid 5 and name H1' and segid 1) 3.3 2.0 2.0
assign (resid 5 and name H2' and segid 1) (resid 5 and name H4' and segid 1) 2.8 1.1 1.1
assign (resid 5 and name H2" and segid 1) (resid 5 and name H4' and segid 1) 2.3 1.5 1.5
assign (resid 5 and name H5' and segid 1) (resid 5 and name H1' and segid 1) 3.3 1.3 1.3
assign (resid 5 and name H1' and segid 1) (resid 5 and name H8 and segid 1) 1.9 1.2 1.2
assign (resid 5 and name H2' and segid 1) (resid 5 and name H2" and segid 1) 2.4 1.0 1.0
assign (resid 5 and name H2' and segid 1) (resid 5 and name H8 and segid 1) 3.2 1.3 1.3
assign (resid 5 and name H3' and segid 1) (resid 5 and name H1' and segid 1) 2.9 1.2 1.2
assign (resid 5 and name H4' and segid 1) (resid 5 and name H8 and segid 1) 3.5 2.0 2.0
assign (resid 5 and name H8 and segid 1) (resid 6 and name H8 and segid 1) 3.8 1.5 1.5
assign (resid 6 and name H2" and segid 1) (resid 5 and name H8 and segid 1) 3.0 2.0 2.0
assign (resid 6 and name H1' and segid 1) (resid 5 and name H8 and segid 1) 5.6 2.2 2.2
assign (resid 6 and name H2' and segid 1) (resid 5 and name H8 and segid 1) 2.3 1.5 1.5
assign (resid 6 and name H3' and segid 1) (resid 5 and name H8 and segid 1) 3.7 1.5 1.5
assign (resid 6 and name H2' and segid 1) (resid 6 and name H2" and segid 1) 2.7 1.1 1.1
assign (resid 6 and name H2' and segid 1) (resid 6 and name H3' and segid 1) 3.3 1.3 1.3
assign (resid 6 and name H2' and segid 1) (resid 6 and name H8 and segid 1) 3.0 1.2 1.2
assign (resid 6 and name H2" and segid 1) (resid 6 and name H1' and segid 1) 3.3 1.3 1.3
assign (resid 6 and name H2" and segid 1) (resid 6 and name H8 and segid 1) 2.5 1.0 1.0
assign (resid 6 and name H4' and segid 1) (resid 6 and name H3' and segid 1) 2.7 1.1 1.1
assign (resid 6 and name H2' and segid 1) (resid 6 and name H4' and segid 1) 3.5 1.4 1.4
assign (resid 6 and name H3' and segid 1) (resid 6 and name H8 and segid 1) 3.1 2.0 2.0
assign (resid 6 and name H4' and segid 1) (resid 6 and name H1' and segid 1) 2.8 1.1 1.1
assign (resid 6 and name H1' and segid 1) (resid 6 and name H8 and segid 1) 2.6 1.0 1.0
assign (resid 6 and name H3' and segid 1) (resid 6 and name H1' and segid 1) 3.4 1.4 1.4
assign (resid 6 and name H4' and segid 1) (resid 6 and name H8 and segid 1) 3.7 2.0 2.0
assign (resid 6 and name H5' and segid 1) (resid 6 and name H1' and segid 1) 3.5 1.4 1.4
assign (resid 6 and name H8 and segid 1) (resid 7 and name H8 and segid 1) 4.0 2.0 2.0
assign (resid 7 and name H1' and segid 1) (resid 6 and name H8 and segid 1) 3.7 3.0 3.0
assign (resid 7 and name H2' and segid 1) (resid 7 and name H2" and segid 1) 2.1 0.8 0.8
assign (resid 7 and name H2" and segid 1) (resid 7 and name H1' and segid 1) 2.0 0.8 0.8
assign (resid 7 and name H2" and segid 1) (resid 7 and name H8 and segid 1) 2.6 1.5 1.5
assign (resid 7 and name H4' and segid 1) (resid 7 and name H1' and segid 1) 4 1.6 1.6
assign (resid 7 and name H4' and segid 1) (resid 7 and name H3' and segid 1) 2.5 1.0 1.0
assign (resid 7 and name H5' and segid 1) (resid 7 and name H1' and segid 1) 4.2 1.7 1.7
assign (resid 7 and name H5' and segid 1) (resid 7 and name H3' and segid 1) 2.7 1.1 1.1
assign (resid 7 and name H5' and segid 1) (resid 7 and name H4' and segid 1) 2.4 1.0 1.0
assign (resid 7 and name H3' and segid 1) (resid 7 and name H8 and segid 1) 3.7 2.0 2.0
assign (resid 7 and name H5" and segid 1) (resid 7 and name H3' and segid 1) 2.9 1.2 1.2
assign (resid 7 and name H3' and segid 1) (resid 7 and name H1' and segid 1) 3.1 1.2 1.2
assign (resid 7 and name H5' and segid 1) (resid 7 and name H8 and segid 1) 3.7 2.5 2.5
assign (resid 7 and name H2' and segid 1) (resid 7 and name H1' and segid 1) 2.7 1.1 1.1
assign (resid 7 and name H2' and segid 1) (resid 7 and name H8 and segid 1) 2.6 1.5 1.5

assign (resid 7 and name H4' and segid 1) (resid 7 and name H8 and segid 1) 4.6 1.8 1.8
assign (resid 7 and name H3' and segid 1) (resid 7 and name H1' and segid 1) 4.7 1.9 1.9
assign (resid 8 and name H2' and segid 1) (resid 8 and name H1' and segid 1) 2.7 1.5 1.5
assign (resid 8 and name H2' and segid 1) (resid 8 and name H2" and segid 1) 2.9 1.2 1.2
assign (resid 8 and name H2' and segid 1) (resid 8 and name H8 and segid 1) 2.5 1.5 1.5
assign (resid 8 and name H2" and segid 1) (resid 8 and name H8 and segid 1) 3.1 1.2 1.2
assign (resid 8 and name H1' and segid 1) (resid 8 and name H8 and segid 1) 3.3 1.3 1.3
assign (resid 8 and name H2" and segid 1) (resid 8 and name H1' and segid 1) 1.4 1.0 1.0
assign (resid 8 and name H3' and segid 1) (resid 8 and name H1' and segid 1) 3.1 1.2 1.2
assign (resid 8 and name H3' and segid 1) (resid 8 and name H8 and segid 1) 3.5 2.0 2.0
assign (resid 8 and name H4' and segid 1) (resid 8 and name H8 and segid 1) 3.6 2.0 2.0
assign (resid 9 and name H8 and segid 1) (resid 8 and name H8 and segid 1) 3.3 1.3 1.3
assign (resid 9 and name H2" and segid 1) (resid 8 and name H8 and segid 1) 3.2 2.5 2.5
assign (resid 9 and name H1' and segid 1) (resid 8 and name H8 and segid 1) 3.2 2.5 2.5
assign (resid 9 and name H2' and segid 1) (resid 8 and name H8 and segid 1) 3.2 2.0 2.0
assign (resid 9 and name H2' and segid 1) (resid 9 and name H1' and segid 1) 3.3 1.3 1.3
assign (resid 9 and name H2' and segid 1) (resid 9 and name H2" and segid 1) 2.0 0.8 0.8
assign (resid 9 and name H2" and segid 1) (resid 9 and name H1' and segid 1) 2.7 1.1 1.1
assign (resid 9 and name H4' and segid 1) (resid 9 and name H1' and segid 1) 3.2 1.3 1.3
assign (resid 9 and name H4' and segid 1) (resid 9 and name H3' and segid 1) 2.5 1.0 1.0
assign (resid 9 and name H5" and segid 1) (resid 9 and name H1' and segid 1) 3.1 2.0 2.0
assign (resid 9 and name H2" and segid 1) (resid 9 and name H8 and segid 1) 2.7 1.5 1.5
assign (resid 9 and name H3' and segid 1) (resid 9 and name H1' and segid 1) 3.5 1.4 1.4
assign (resid 9 and name H5' and segid 1) (resid 9 and name H1' and segid 1) 2.9 2.0 2.0
assign (resid 9 and name H1' and segid 1) (resid 9 and name H8 and segid 1) 3.1 1.2 1.2
assign (resid 9 and name H3' and segid 1) (resid 9 and name H8 and segid 1) 4.2 1.7 1.7
assign (resid 9 and name H4' and segid 1) (resid 9 and name H8 and segid 1) 4.2 1.7 1.7
assign (resid 10 and name H2" and segid 1) (resid 10 and name H8 and segid 1) 2.9 1.2 1.2
assign (resid 10 and name H1' and segid 1) (resid 10 and name H8 and segid 1) 3.4 1.4 1.4
assign (resid 10 and name H2' and segid 1) (resid 10 and name H3' and segid 1) 2.2 0.9 0.9
assign (resid 10 and name H2" and segid 1) (resid 10 and name H1' and segid 1) 2.4 1.0 1.0
assign (resid 10 and name H2" and segid 1) (resid 10 and name H3' and segid 1) 2.1 0.8 0.8
assign (resid 10 and name H2" and segid 1) (resid 10 and name H4' and segid 1) 2.6 1.5 1.5
assign (resid 10 and name H4' and segid 1) (resid 10 and name H1' and segid 1) 3.2 1.3 1.3
assign (resid 10 and name H2' and segid 1) (resid 10 and name H1' and segid 1) 2.4 1.0 1.0
assign (resid 10 and name H2' and segid 1) (resid 10 and name H4' and segid 1) 2.7 1.1 1.1
assign (resid 10 and name H3' and segid 1) (resid 10 and name H8 and segid 1) 3.9 1.6 1.6
assign (resid 10 and name H3' and segid 1) (resid 10 and name H1' and segid 1) 4.3 1.7 1.7
assign (resid 10 and name H4' and segid 1) (resid 10 and name H8 and segid 1) 4.0 1.6 1.6
assign (resid 11 and name H8 and segid 1) (resid 10 and name H8 and segid 1) 4.0 1.6 1.6
assign (resid 11 and name H3' and segid 1) (resid 10 and name H8 and segid 1) 4.6 1.8 1.8
assign (resid 11 and name H1' and segid 1) (resid 10 and name H8 and segid 1) 3.5 2.0 2.0
assign (resid 11 and name H2" and segid 1) (resid 10 and name H8 and segid 1) 3.1 2.5 2.5
assign (resid 11 and name H2' and segid 1) (resid 10 and name H8 and segid 1) 3.2 2.0 2.0
assign (resid 11 and name H2' and segid 1) (resid 11 and name H2" and segid 1) 2.3 0.9 0.9
assign (resid 11 and name H2' and segid 1) (resid 11 and name H3' and segid 1) 3.1 1.2 1.2
assign (resid 11 and name H2' and segid 1) (resid 11 and name H5' and segid 1) 2.3 1.5 1.5
assign (resid 11 and name H2' and segid 1) (resid 11 and name H8 and segid 1) 2.6 1.0 1.0
assign (resid 11 and name H2" and segid 1) (resid 11 and name H1' and segid 1) 2.7 1.1 1.1
assign (resid 11 and name H2" and segid 1) (resid 11 and name H3' and segid 1) 2.6 1.0 1.0
assign (resid 11 and name H2" and segid 1) (resid 11 and name H4' and segid 1) 4.1 1.6 1.6
assign (resid 11 and name H2" and segid 1) (resid 11 and name H8 and segid 1) 2.9 1.2 1.2
assign (resid 11 and name H5" and segid 1) (resid 11 and name H4' and segid 1) 2.9 1.2 1.2
assign (resid 11 and name H1' and segid 1) (resid 11 and name H8 and segid 1) 3.3 1.3 1.3
assign (resid 11 and name H2' and segid 1) (resid 11 and name H4' and segid 1) 2.7 1.1 1.1
assign (resid 11 and name H3' and segid 1) (resid 11 and name H8 and segid 1) 3.4 2.0 2.0
assign (resid 11 and name H2' and segid 1) (resid 11 and name H1' and segid 1) 4.0 1.6 1.6
assign (resid 11 and name H2" and segid 1) (resid 11 and name H5' and segid 1) 3.6 1.4 1.4
assign (resid 11 and name H3' and segid 1) (resid 11 and name H1' and segid 1) 3.4 1.4 1.4
assign (resid 11 and name H4' and segid 1) (resid 11 and name H8 and segid 1) 3.7 2.0 2.0

assign (resid 11 and name H5' and segid 1) (resid 11 and name H8 and segid 1) 3.4 2.5 2.5
assign (resid 12 and name H2' and segid 1) (resid 11 and name H8 and segid 1) 3.4 1.7 1.7
assign (resid 12 and name H2" and segid 1) (resid 11 and name H8 and segid 1) 3.2 1.6 1.6
assign (resid 12 and name H2' and segid 1) (resid 12 and name H5' and segid 1) 2.1 1.2 1.2
assign (resid 12 and name H2" and segid 1) (resid 12 and name H1' and segid 1) 2.8 1.4 1.4
assign (resid 12 and name H5' and segid 1) (resid 12 and name H1' and segid 1) 2.9 1.5 1.5
assign (resid 12 and name H5" and segid 1) (resid 12 and name H1' and segid 1) 3.2 2.0 2.0
assign (resid 12 and name H2' and segid 1) (resid 12 and name H8 and segid 1) 3.1 1.6 1.6
assign (resid 12 and name H2" and segid 1) (resid 12 and name H8 and segid 1) 3.5 1.8 1.8
assign (resid 12 and name H2' and segid 1) (resid 12 and name H4' and segid 1) 4.1 2.1 2.1
assign (resid 12 and name H1' and segid 1) (resid 12 and name H8 and segid 1) 3.8 1.9 1.9
assign (resid 12 and name H2' and segid 1) (resid 12 and name H2" and segid 1) 2.7 1.4 1.4
assign (resid 12 and name H4' and segid 1) (resid 12 and name H1' and segid 1) 4.0 2.0 2.0
assign (resid 14 and name H1' and segid 1) (resid 14 and name H6 and segid 1) 2.2 1.5 1.5
assign (resid 14 and name H2' and segid 1) (resid 14 and name H5" and segid 1) 3.1 1.6 1.6
assign (resid 14 and name H2" and segid 1) (resid 14 and name H1' and segid 1) 3.7 1.7 1.7
assign (resid 14 and name H2" and segid 1) (resid 14 and name H4' and segid 1) 2.6 1.3 1.3
assign (resid 14 and name H5 and segid 1) (resid 14 and name H6 and segid 1) 2.4 1.2 1.2
assign (resid 14 and name H2' and segid 1) (resid 14 and name H2" and segid 1) 1.6 0.8 0.8
assign (resid 14 and name H2' and segid 1) (resid 14 and name H6 and segid 1) 1.8 0.9 0.9
assign (resid 14 and name H4' and segid 1) (resid 14 and name H6 and segid 1) 3.3 2.0 2.0
assign (resid 14 and name H5" and segid 1) (resid 14 and name H6 and segid 1) 3.3 1.7 1.7
assign (resid 14 and name H2' and segid 1) (resid 14 and name H5' and segid 1) 3.3 1.7 1.7
assign (resid 14 and name H2" and segid 1) (resid 14 and name H5' and segid 1) 3.3 1.7 1.7
assign (resid 14 and name H2' and segid 1) (resid 14 and name H1' and segid 1) 4.0 2.0 2.0
assign (resid 14 and name H2" and segid 1) (resid 14 and name H5" and segid 1) 2.6 2.0 2.0
assign (resid 14 and name H5" and segid 1) (resid 15 and name H6 and segid 1) 3.6 1.7 1.7
assign (resid 15 and name H1' and segid 1) (resid 15 and name H6 and segid 1) 2.6 1.0 1.0
assign (resid 15 and name H2' and segid 1) (resid 15 and name H1' and segid 1) 2.6 1.0 1.0
assign (resid 15 and name H2' and segid 1) (resid 15 and name H2" and segid 1) 2.1 0.8 0.8
assign (resid 15 and name H2" and segid 1) (resid 15 and name H1' and segid 1) 1.4 1.0 1.0
assign (resid 15 and name H4' and segid 1) (resid 15 and name H6 and segid 1) 3.6 1.4 1.4
assign (resid 15 and name H5 and segid 1) (resid 15 and name H6 and segid 1) 2.4 1.0 1.0
assign (resid 15 and name H5" and segid 1) (resid 15 and name H6 and segid 1) 3.9 1.6 1.6
assign (resid 15 and name H5' and segid 1) (resid 15 and name H6 and segid 1) 3.5 1.4 1.4
assign (resid 15 and name H2' and segid 1) (resid 15 and name H6 and segid 1) 2.6 1.5 1.5
assign (resid 15 and name H2" and segid 1) (resid 15 and name H6 and segid 1) 2.7 2.0 2.0
assign (resid 15 and name H2" and segid 1) (resid 16 and name H6 and segid 1) 2.4 1.0 1.0
assign (resid 16 and name H7* and segid 1) (resid 15 and name H1' and segid 1) 2.8 1.1 1.1
assign (resid 16 and name H7* and segid 1) (resid 15 and name H6 and segid 1) 2.8 1.1 1.1
assign (resid 16 and name H7* and segid 1) (resid 15 and name H5 and segid 1) 3.0 2.0 2.0
assign (resid 16 and name H1' and segid 1) (resid 16 and name H6 and segid 1) 3.8 1.5 1.5
assign (resid 16 and name H2' and segid 1) (resid 16 and name H2" and segid 1) 3.0 1.2 1.2
assign (resid 16 and name H2' and segid 1) (resid 16 and name H6 and segid 1) 1.4 1.0 1.0
assign (resid 16 and name H2" and segid 1) (resid 16 and name H4' and segid 1) 3.2 1.3 1.3
assign (resid 16 and name H2" and segid 1) (resid 16 and name H6 and segid 1) 2.9 1.2 1.2
assign (resid 16 and name H7* and segid 1) (resid 16 and name H1' and segid 1) 3.1 2.0 2.0
assign (resid 16 and name H7* and segid 1) (resid 16 and name H6 and segid 1) 2.1 0.8 0.8
assign (resid 16 and name H2' and segid 1) (resid 16 and name H1' and segid 1) 1.9 1.0 1.0
assign (resid 16 and name H2' and segid 1) (resid 16 and name H5" and segid 1) 3.1 2.0 2.0
assign (resid 16 and name H4' and segid 1) (resid 16 and name H1' and segid 1) 2.2 1.5 1.5
assign (resid 16 and name H2" and segid 1) (resid 16 and name H1' and segid 1) 3.1 1.2 1.2
assign (resid 16 and name H2" and segid 1) (resid 17 and name H6 and segid 1) 3.8 1.5 1.5
assign (resid 17 and name H7* and segid 1) (resid 17 and name H6 and segid 1) 2.7 1.1 1.1
assign (resid 17 and name H2' and segid 1) (resid 17 and name H1' and segid 1) 2.8 1.1 1.1
assign (resid 17 and name H2' and segid 1) (resid 17 and name H2" and segid 1) 1.7 0.7 0.7
assign (resid 17 and name H1' and segid 1) (resid 17 and name H6 and segid 1) 2.1 1.5 1.5
assign (resid 17 and name H2" and segid 1) (resid 17 and name H1' and segid 1) 2.6 1.0 1.0
assign (resid 17 and name H7* and segid 1) (resid 18 and name H5 and segid 1) 3.4 1.4 1.4
assign (resid 18 and name H2' and segid 1) (resid 18 and name H2" and segid 1) 2.8 1.1 1.1

assign (resid 18 and name H1' and segid 1) (resid 18 and name H6 and segid 1) 3.9 1.6 1.6
assign (resid 18 and name H2' and segid 1) (resid 18 and name H6 and segid 1) 2.3 1.5 1.5
assign (resid 18 and name H2" and segid 1) (resid 18 and name H6 and segid 1) 4.0 1.6 1.6
assign (resid 18 and name H5 and segid 1) (resid 18 and name H6 and segid 1) 2.5 1.5 1.5
assign (resid 19 and name H2" and segid 1) (resid 19 and name H6 and segid 1) 2.4 1.5 1.5
assign (resid 19 and name H2" and segid 1) (resid 19 and name H1' and segid 1) 2.7 1.1 1.1
assign (resid 19 and name H5 and segid 1) (resid 19 and name H6 and segid 1) 2.6 1.0 1.0
assign (resid 19 and name H2' and segid 1) (resid 19 and name H1' and segid 1) 2.7 1.1 1.1
assign (resid 19 and name H2' and segid 1) (resid 19 and name H6 and segid 1) 2.1 2.0 2.0
assign (resid 20 and name H7* and segid 1) (resid 19 and name H6 and segid 1) 2.8 1.1 1.1
assign (resid 20 and name H7* and segid 1) (resid 19 and name H5 and segid 1) 3.0 2.0 2.0
assign (resid 20 and name H1' and segid 1) (resid 20 and name H6 and segid 1) 2.8 1.1 1.1
assign (resid 20 and name H2' and segid 1) (resid 20 and name H1' and segid 1) 2.4 1.0 1.0
assign (resid 20 and name H2' and segid 1) (resid 20 and name H2" and segid 1) 2.1 0.8 0.8
assign (resid 20 and name H2" and segid 1) (resid 20 and name H1' and segid 1) 1.8 0.7 0.7
assign (resid 20 and name H2' and segid 1) (resid 20 and name H6 and segid 1) 2.5 1.5 1.5
assign (resid 20 and name H3' and segid 1) (resid 20 and name H1' and segid 1) 2.4 2.0 2.0
assign (resid 20 and name H7* and segid 1) (resid 20 and name H1' and segid 1) 2.8 2.5 2.5
assign (resid 20 and name H7* and segid 1) (resid 20 and name H6 and segid 1) 2.3 0.9 0.9
assign (resid 21 and name H2' and segid 1) (resid 21 and name H2" and segid 1) 3.0 1.2 1.2
assign (resid 21 and name H5' and segid 1) (resid 21 and name H6 and segid 1) 3.2 2.0 2.0
assign (resid 21 and name H1' and segid 1) (resid 21 and name H6 and segid 1) 3.6 1.4 1.4
assign (resid 21 and name H2' and segid 1) (resid 21 and name H6 and segid 1) 2.2 1.5 1.5
assign (resid 21 and name H2" and segid 1) (resid 21 and name H1' and segid 1) 2.3 0.9 0.9
assign (resid 21 and name H7* and segid 1) (resid 21 and name H1' and segid 1) 3.1 2.0 2.0
assign (resid 21 and name H7* and segid 1) (resid 21 and name H6 and segid 1) 2.5 1.0 1.0
assign (resid 21 and name H2' and segid 1) (resid 21 and name H1' and segid 1) 2.1 0.8 0.8
assign (resid 21 and name H2" and segid 1) (resid 21 and name H6 and segid 1) 3.0 1.2 1.2
assign (resid 21 and name H5' and segid 1) (resid 21 and name H4' and segid 1) 1.6 1.0 1.0
assign (resid 22 and name H5 and segid 1) (resid 22 and name H6 and segid 1) 2.6 1.0 1.0
assign (resid 21 and name H7* and segid 1) (resid 22 and name H5 and segid 1) 3.0 2.5 2.5
assign (resid 22 and name H1' and segid 1) (resid 22 and name H6 and segid 1) 2.9 1.2 1.2
assign (resid 22 and name H2" and segid 1) (resid 22 and name H5 and segid 1) 3.1 2.0 2.0
assign (resid 22 and name H4' and segid 1) (resid 22 and name H6 and segid 1) 3.3 2.5 2.5
assign (resid 22 and name H2' and segid 1) (resid 22 and name H5 and segid 1) 2.7 2.0 2.0
assign (resid 22 and name H2' and segid 1) (resid 22 and name H6 and segid 1) 2.8 1.1 1.1
assign (resid 22 and name H2" and segid 1) (resid 22 and name H6 and segid 1) 2.2 1.5 1.5
assign (resid 23 and name H2' and segid 1) (resid 23 and name H1' and segid 1) 2.7 1.1 1.1
assign (resid 23 and name H2' and segid 1) (resid 23 and name H2" and segid 1) 2.3 0.9 0.9
assign (resid 23 and name H2' and segid 1) (resid 23 and name H6 and segid 1) 2.0 1.5 1.5
assign (resid 23 and name H2" and segid 1) (resid 23 and name H1' and segid 1) 2.6 1.0 1.0
assign (resid 23 and name H5 and segid 1) (resid 23 and name H6 and segid 1) 2.6 1.0 1.0
assign (resid 23 and name H1' and segid 1) (resid 23 and name H6 and segid 1) 2.7 1.1 1.1
assign (resid 23 and name H2" and segid 1) (resid 23 and name H6 and segid 1) 2.2 2.0 2.0
assign (resid 23 and name H5' and segid 1) (resid 23 and name H1' and segid 1) 4.3 1.7 1.7
assign (resid 24 and name H7* and segid 1) (resid 23 and name H1' and segid 1) 3.0 1.2 1.2
assign (resid 24 and name H7* and segid 1) (resid 23 and name H5 and segid 1) 3.0 2.0 2.0
assign (resid 24 and name H7* and segid 1) (resid 23 and name H6 and segid 1) 2.1 2.0 2.0
assign (resid 24 and name H4' and segid 1) (resid 24 and name H6 and segid 1) 4.6 1.8 1.8
assign (resid 24 and name H1' and segid 1) (resid 24 and name H6 and segid 1) 2.9 1.2 1.2
assign (resid 24 and name H4' and segid 1) (resid 24 and name H1' and segid 1) 4.9 2.0 2.0
assign (resid 24 and name H7* and segid 1) (resid 24 and name H6 and segid 1) 2.1 0.8 0.8
assign (resid 24 and name H2' and segid 1) (resid 24 and name H6 and segid 1) 1.9 1.5 1.5
assign (resid 24 and name H2" and segid 1) (resid 24 and name H6 and segid 1) 2.9 1.2 1.2
assign (resid 24 and name H3' and segid 1) (resid 24 and name H1' and segid 1) 3.5 1.4 1.4
assign (resid 24 and name H5' and segid 1) (resid 24 and name H6 and segid 1) 2.7 2.0 2.0
assign (resid 24 and name H2' and segid 1) (resid 25 and name H6 and segid 1) 3.0 1.5 1.5
assign (resid 25 and name H2" and segid 1) (resid 25 and name H6 and segid 1) 2.3 2.0 2.0
assign (resid 25 and name H4' and segid 1) (resid 25 and name H6 and segid 1) 4.1 2.1 2.1
assign (resid 25 and name H5 and segid 1) (resid 25 and name H6 and segid 1) 2.5 1.3 1.3

assign (resid 25 and name H5' and segid 1) (resid 25 and name H4' and segid 1) 1.5 1.2 1.2
assign (resid 25 and name H2' and segid 1) (resid 25 and name H1' and segid 1) 2.3 1.2 1.2
assign (resid 25 and name H3' and segid 1) (resid 25 and name H1' and segid 1) 3.5 1.8 1.8
assign (resid 25 and name H4' and segid 1) (resid 25 and name H3' and segid 1) 3.0 1.5 1.5
assign (resid 25 and name H5' and segid 1) (resid 25 and name H6 and segid 1) 4.2 2.1 2.1
assign (resid 25 and name H1' and segid 1) (resid 25 and name H6 and segid 1) 3.1 1.6 1.6
assign (resid 25 and name H2' and segid 1) (resid 25 and name H3' and segid 1) 2.6 1.3 1.3
assign (resid 25 and name H2' and segid 1) (resid 25 and name H5' and segid 1) 2.9 2.5 2.5
assign (resid 25 and name H2' and segid 1) (resid 25 and name H6 and segid 1) 2.4 1.2 1.2
assign (resid 25 and name H3' and segid 1) (resid 25 and name H6 and segid 1) 3.0 1.5 1.5
assign (resid 25 and name H5' and segid 1) (resid 25 and name H1' and segid 1) 2.9 2.0 2.0
assign (resid 25 and name H5' and segid 1) (resid 25 and name H3' and segid 1) 2.9 1.5 1.5
assign (resid 25 and name H4' and segid 1) (resid 25 and name H1' and segid 1) 3.7 1.9 1.9
assign (resid 25 and name H5" and segid 1) (resid 25 and name H4' and segid 1) 1.9 1.0 1.0

Build up non-exchangeable proton distance restraints file

assign (resid 1 and name H2' and segid 1) (resid 1 and name H8 and segid 1) 3.2 1.9 1.9
assign (resid 1 and name H2" and segid 1) (resid 1 and name H4' and segid 1) 2.5 1.5 1.5
assign (resid 1 and name H2" and segid 1) (resid 1 and name H1' and segid 1) 1.9 1.1 1.1
assign (resid 1 and name H1' and segid 1) (resid 1 and name H8 and segid 1) 1.8 1.8 1.8
assign (resid 1 and name H2' and segid 1) (resid 1 and name H2" and segid 1) 1.7 1.0 1.0
assign (resid 1 and name H2" and segid 1) (resid 1 and name H8 and segid 1) 3.0 1.8 1.8
assign (resid 1 and name H4' and segid 1) (resid 1 and name H1' and segid 1) 3.8 2.3 2.3
assign (resid 1 and name H4' and segid 1) (resid 1 and name H8 and segid 1) 3.2 2.5 2.5
assign (resid 1 and name H5' and segid 1) (resid 1 and name H8 and segid 1) 3.6 2.1 2.1
assign (resid 2 and name H1' and segid 1) (resid 1 and name H8 and segid 1) 3.1 3.0 3.0
assign (resid 2 and name H1' and segid 1) (resid 2 and name H8 and segid 1) 4.6 4.0 1.8
assign (resid 2 and name H2' and segid 1) (resid 2 and name H1' and segid 1) 2.5 1.0 1.0
assign (resid 2 and name H2' and segid 1) (resid 2 and name H2" and segid 1) 1.8 0.7 0.7
assign (resid 2 and name H2' and segid 1) (resid 2 and name H8 and segid 1) 2.1 2.0 2.0
assign (resid 2 and name H2" and segid 1) (resid 2 and name H1' and segid 1) 2.4 0.9 0.9
assign (resid 2 and name H4' and segid 1) (resid 2 and name H1' and segid 1) 2.9 1.2 1.2
assign (resid 2 and name H5' and segid 1) (resid 2 and name H4' and segid 1) 2.0 0.8 0.8
assign (resid 2 and name H2" and segid 1) (resid 2 and name H4' and segid 1) 3.5 1.4 1.4
assign (resid 2 and name H5" and segid 1) (resid 2 and name H8 and segid 1) 3.0 2.0 3.0
assign (resid 2 and name H2' and segid 1) (resid 2 and name H4' and segid 1) 4.6 1.8 1.8
assign (resid 2 and name H2" and segid 1) (resid 2 and name H8 and segid 1) 3.2 1.3 1.3
assign (resid 2 and name H4' and segid 1) (resid 2 and name H8 and segid 1) 4.8 1.9 1.9
assign (resid 3 and name H1' and segid 1) (resid 2 and name H8 and segid 1) 3.0 2.5 3.0
assign (resid 3 and name H2' and segid 1) (resid 3 and name H1' and segid 1) 3.6 1.4 1.4
assign (resid 3 and name H1' and segid 1) (resid 3 and name H8 and segid 1) 2.8 1.1 1.1
assign (resid 3 and name H2' and segid 1) (resid 3 and name H2" and segid 1) 1.7 0.7 0.7
assign (resid 3 and name H2' and segid 1) (resid 3 and name H8 and segid 1) 3.2 1.3 1.3
assign (resid 3 and name H2" and segid 1) (resid 3 and name H1' and segid 1) 2.1 0.8 0.8
assign (resid 3 and name H2' and segid 1) (resid 3 and name H4' and segid 1) 3.4 1.4 1.4
assign (resid 3 and name H2" and segid 1) (resid 3 and name H8 and segid 1) 3.9 1.6 1.6
assign (resid 3 and name H3' and segid 1) (resid 3 and name H1' and segid 1) 3.4 1.4 1.4
assign (resid 3 and name H4' and segid 1) (resid 3 and name H8 and segid 1) 3.0 2.0 3.0
assign (resid 3 and name H4' and segid 1) (resid 3 and name H1' and segid 1) 2.7 1.1 1.1
assign (resid 3 and name H8 and segid 1) (resid 4 and name H8 and segid 1) 4.0 2.0 2.0
assign (resid 4 and name H2' and segid 1) (resid 4 and name H2" and segid 1) 2.2 0.9 0.9
assign (resid 4 and name H2' and segid 1) (resid 4 and name H3' and segid 1) 2.2 0.9 0.9
assign (resid 4 and name H2' and segid 1) (resid 4 and name H8 and segid 1) 2.2 2.0 2.0
assign (resid 4 and name H2" and segid 1) (resid 4 and name H1' and segid 1) 2.5 1.0 1.0
assign (resid 4 and name H2" and segid 1) (resid 4 and name H3' and segid 1) 2.0 1.0 1.0
assign (resid 4 and name H2" and segid 1) (resid 4 and name H4' and segid 1) 2.4 1.8 1.8

assign (resid 4 and name H5' and segid 1) (resid 4 and name H4' and segid 1) 1.8 1.8 1.8
assign (resid 4 and name H1' and segid 1) (resid 4 and name H8 and segid 1) 3.4 1.4 1.4
assign (resid 4 and name H3' and segid 1) (resid 4 and name H1' and segid 1) 4.2 1.7 1.7
assign (resid 4 and name H5" and segid 1) (resid 4 and name H4' and segid 1) 2.1 0.9 0.9
assign (resid 4 and name H4' and segid 1) (resid 4 and name H8 and segid 1) 2.9 2.5 2.5
assign (resid 4 and name H2' and segid 1) (resid 4 and name H1' and segid 1) 2.5 1.0 1.0
assign (resid 4 and name H2" and segid 1) (resid 4 and name H8 and segid 1) 2.8 2.0 2.0
assign (resid 4 and name H4' and segid 1) (resid 4 and name H1' and segid 1) 3.1 1.2 1.2
assign (resid 4 and name H5" and segid 1) (resid 4 and name H8 and segid 1) 3.7 3.0 3.0
assign (resid 5 and name H8 and segid 1) (resid 4 and name H8 and segid 1) 4.5 2.5 2.5
assign (resid 5 and name H1' and segid 1) (resid 4 and name H8 and segid 1) 3.5 3.0 3.0
assign (resid 5 and name H2' and segid 1) (resid 4 and name H8 and segid 1) 3.2 1.3 1.3
assign (resid 5 and name H2' and segid 1) (resid 5 and name H1' and segid 1) 1.9 1.0 1.0
assign (resid 5 and name H2" and segid 1) (resid 5 and name H1' and segid 1) 2.3 0.9 0.9
assign (resid 5 and name H2" and segid 1) (resid 5 and name H8 and segid 1) 2.4 2.0 2.0
assign (resid 5 and name H4' and segid 1) (resid 5 and name H1' and segid 1) 2.8 1.1 1.1
assign (resid 5 and name H2' and segid 1) (resid 5 and name H4' and segid 1) 3.1 1.3 1.3
assign (resid 5 and name H2" and segid 1) (resid 5 and name H4' and segid 1) 3.0 1.2 1.2
assign (resid 5 and name H5' and segid 1) (resid 5 and name H1' and segid 1) 3.3 1.3 1.3
assign (resid 5 and name H1' and segid 1) (resid 5 and name H8 and segid 1) 1.9 1.5 1.9
assign (resid 5 and name H2' and segid 1) (resid 5 and name H2" and segid 1) 2.4 1.0 1.0
assign (resid 5 and name H2' and segid 1) (resid 5 and name H8 and segid 1) 3.2 2.2 2.2
assign (resid 5 and name H3' and segid 1) (resid 5 and name H1' and segid 1) 2.9 1.2 1.2
assign (resid 5 and name H8 and segid 1) (resid 6 and name H8 and segid 1) 3.8 3.0 3.0
assign (resid 6 and name H1' and segid 1) (resid 5 and name H8 and segid 1) 2.6 2.6 2.6
assign (resid 6 and name H3' and segid 1) (resid 5 and name H8 and segid 1) 3.7 3.0 3.0
assign (resid 6 and name H2' and segid 1) (resid 6 and name H2" and segid 1) 2.1 0.9 0.9
assign (resid 6 and name H2' and segid 1) (resid 6 and name H3' and segid 1) 3.1 1.2 1.2
assign (resid 6 and name H2' and segid 1) (resid 6 and name H8 and segid 1) 1.7 1.5 1.5
assign (resid 6 and name H2" and segid 1) (resid 6 and name H1' and segid 1) 3.3 1.3 1.3
assign (resid 6 and name H2" and segid 1) (resid 6 and name H8 and segid 1) 2.4 2.0 2.0
assign (resid 6 and name H4' and segid 1) (resid 6 and name H3' and segid 1) 2.7 1.1 1.1
assign (resid 6 and name H2' and segid 1) (resid 6 and name H4' and segid 1) 3.5 1.4 1.4
assign (resid 6 and name H2" and segid 1) (resid 6 and name H1' and segid 1) 1.7 0.7 0.7
assign (resid 6 and name H3' and segid 1) (resid 6 and name H8 and segid 1) 3.0 1.8 1.8
assign (resid 6 and name H4' and segid 1) (resid 6 and name H1' and segid 1) 2.8 1.1 1.1
assign (resid 6 and name H1' and segid 1) (resid 6 and name H8 and segid 1) 3.2 1.3 1.3
assign (resid 6 and name H3' and segid 1) (resid 6 and name H1' and segid 1) 2.9 1.2 1.2
assign (resid 6 and name H4' and segid 1) (resid 6 and name H8 and segid 1) 3.9 1.6 1.6
assign (resid 6 and name H5' and segid 1) (resid 6 and name H1' and segid 1) 3.5 1.4 1.4
assign (resid 6 and name H8 and segid 1) (resid 7 and name H8 and segid 1) 4.0 2.5 2.5
assign (resid 7 and name H2' and segid 1) (resid 7 and name H2" and segid 1) 1.7 0.7 0.7
assign (resid 7 and name H2" and segid 1) (resid 7 and name H1' and segid 1) 3.6 1.4 1.4
assign (resid 7 and name H4' and segid 1) (resid 7 and name H1' and segid 1) 2.7 1.1 1.1
assign (resid 7 and name H4' and segid 1) (resid 7 and name H3' and segid 1) 2.2 0.9 0.9
assign (resid 7 and name H5' and segid 1) (resid 7 and name H1' and segid 1) 3.5 1.4 1.4
assign (resid 7 and name H5' and segid 1) (resid 7 and name H3' and segid 1) 2.3 1.5 1.5
assign (resid 7 and name H5' and segid 1) (resid 7 and name H4' and segid 1) 2.0 0.8 0.8
assign (resid 7 and name H3' and segid 1) (resid 7 and name H8 and segid 1) 2.6 2.6 3.0
assign (resid 7 and name H5" and segid 1) (resid 7 and name H3' and segid 1) 3.1 1.2 1.2
assign (resid 7 and name H3' and segid 1) (resid 7 and name H1' and segid 1) 3.1 1.2 1.2
assign (resid 7 and name H2' and segid 1) (resid 7 and name H1' and segid 1) 2.8 1.1 1.1
assign (resid 7 and name H2' and segid 1) (resid 7 and name H8 and segid 1) 2.6 1.5 1.5
assign (resid 7 and name H4' and segid 1) (resid 7 and name H8 and segid 1) 4.6 1.8 1.8
assign (resid 8 and name H1' and segid 1) (resid 7 and name H8 and segid 1) 3.1 2.5 2.5
assign (resid 8 and name H2' and segid 1) (resid 8 and name H1' and segid 1) 2.7 1.1 1.1
assign (resid 8 and name H2' and segid 1) (resid 8 and name H2" and segid 1) 2.5 1.0 1.0
assign (resid 8 and name H2' and segid 1) (resid 8 and name H8 and segid 1) 2.2 1.5 1.5
assign (resid 8 and name H2" and segid 1) (resid 8 and name H8 and segid 1) 2.8 1.5 1.5
assign (resid 8 and name H1' and segid 1) (resid 8 and name H8 and segid 1) 3.0 1.2 1.2

assign (resid 8 and name H5' and segid 1) (resid 8 and name H8 and segid 1) 3.4 2.5 2.5
assign (resid 8 and name H2" and segid 1) (resid 8 and name H1' and segid 1) 2.9 1.1 1.1
assign (resid 8 and name H3' and segid 1) (resid 8 and name H1' and segid 1) 3.1 1.2 1.2
assign (resid 8 and name H3' and segid 1) (resid 8 and name H8 and segid 1) 3.5 2.5 2.5
assign (resid 8 and name H4' and segid 1) (resid 8 and name H8 and segid 1) 3.6 2.0 2.5
assign (resid 8 and name H5" and segid 1) (resid 8 and name H8 and segid 1) 3.4 2.5 3.0
assign (resid 9 and name H2" and segid 1) (resid 8 and name H8 and segid 1) 3.2 3.2 3.2
assign (resid 9 and name H2' and segid 1) (resid 8 and name H8 and segid 1) 4.5 2.5 2.5
assign (resid 9 and name H2' and segid 1) (resid 9 and name H1' and segid 1) 2.8 1.1 1.1
assign (resid 9 and name H2' and segid 1) (resid 9 and name H2" and segid 1) 1.6 0.7 0.7
assign (resid 9 and name H2" and segid 1) (resid 9 and name H1' and segid 1) 2.4 0.9 0.9
assign (resid 9 and name H4' and segid 1) (resid 9 and name H1' and segid 1) 3.2 1.3 1.3
assign (resid 9 and name H4' and segid 1) (resid 9 and name H3' and segid 1) 2.3 0.9 0.9
assign (resid 9 and name H2" and segid 1) (resid 9 and name H8 and segid 1) 2.5 2.0 2.5
assign (resid 9 and name H3' and segid 1) (resid 9 and name H1' and segid 1) 3.0 1.2 1.2
assign (resid 9 and name H1' and segid 1) (resid 9 and name H8 and segid 1) 3.1 1.3 1.3
assign (resid 9 and name H3' and segid 1) (resid 9 and name H8 and segid 1) 4.2 1.7 1.7
assign (resid 9 and name H4' and segid 1) (resid 9 and name H8 and segid 1) 4.2 1.7 1.7
assign (resid 10 and name H1' and segid 1) (resid 10 and name H8 and segid 1) 4.1 2.0 2.0
assign (resid 10 and name H2' and segid 1) (resid 10 and name H3' and segid 1) 2.1 0.8 0.8
assign (resid 10 and name H2" and segid 1) (resid 10 and name H1' and segid 1) 2.2 0.9 0.9
assign (resid 10 and name H2" and segid 1) (resid 10 and name H3' and segid 1) 2.1 0.8 0.8
assign (resid 10 and name H2" and segid 1) (resid 10 and name H4' and segid 1) 2.7 1.1 1.1
assign (resid 10 and name H4' and segid 1) (resid 10 and name H1' and segid 1) 2.8 1.1 1.1
assign (resid 10 and name H1' and segid 1) (resid 9 and name H8 and segid 1) 3.3 2.5 3.0
assign (resid 10 and name H2' and segid 1) (resid 10 and name H1' and segid 1) 2.3 0.9 0.9
assign (resid 10 and name H2' and segid 1) (resid 10 and name H4' and segid 1) 2.9 1.1 1.1
assign (resid 10 and name H3' and segid 1) (resid 10 and name H8 and segid 1) 3.9 2.5 2.5
assign (resid 10 and name H3' and segid 1) (resid 10 and name H1' and segid 1) 4.3 1.7 1.7
assign (resid 10 and name H4' and segid 1) (resid 10 and name H8 and segid 1) 4.0 1.6 1.6
assign (resid 11 and name H8 and segid 1) (resid 10 and name H8 and segid 1) 4.0 2.5 2.5
assign (resid 11 and name H3' and segid 1) (resid 10 and name H8 and segid 1) 4.6 1.9 1.9
assign (resid 11 and name H1' and segid 1) (resid 10 and name H8 and segid 1) 4.0 1.6 1.6
assign (resid 11 and name H2' and segid 1) (resid 11 and name H2" and segid 1) 2.0 0.8 0.8
assign (resid 11 and name H2' and segid 1) (resid 11 and name H3' and segid 1) 2.6 1.1 1.1
assign (resid 11 and name H2' and segid 1) (resid 11 and name H8 and segid 1) 2.3 1.5 1.5
assign (resid 11 and name H2" and segid 1) (resid 10 and name H8 and segid 1) 3.0 2.5 3.0
assign (resid 11 and name H2' and segid 1) (resid 10 and name H8 and segid 1) 3.4 1.5 1.5
assign (resid 11 and name H2" and segid 1) (resid 11 and name H1' and segid 1) 2.4 0.9 0.9
assign (resid 11 and name H2" and segid 1) (resid 11 and name H3' and segid 1) 2.3 0.9 0.9
assign (resid 11 and name H5" and segid 1) (resid 11 and name H3' and segid 1) 3.6 1.5 1.5
assign (resid 11 and name H5" and segid 1) (resid 11 and name H4' and segid 1) 2.5 1.0 1.0
assign (resid 11 and name H1' and segid 1) (resid 11 and name H8 and segid 1) 3.3 1.3 1.3
assign (resid 11 and name H2' and segid 1) (resid 11 and name H4' and segid 1) 3.3 1.3 1.3
assign (resid 11 and name H3' and segid 1) (resid 11 and name H8 and segid 1) 3.3 2.0 2.0
assign (resid 11 and name H2' and segid 1) (resid 11 and name H1' and segid 1) 3.6 1.4 1.4
assign (resid 11 and name H2" and segid 1) (resid 11 and name H5' and segid 1) 3.6 3.0 3.0
assign (resid 11 and name H3' and segid 1) (resid 11 and name H1' and segid 1) 3.5 1.4 1.4
assign (resid 11 and name H4' and segid 1) (resid 11 and name H8 and segid 1) 3.7 2.0 3.0
assign (resid 11 and name H5' and segid 1) (resid 11 and name H8 and segid 1) 3.4 2.0 3.0
assign (resid 12 and name H2" and segid 1) (resid 12 and name H1' and segid 1) 2.4 1.0 1.0
assign (resid 12 and name H5' and segid 1) (resid 12 and name H1' and segid 1) 2.9 2.0 2.0
assign (resid 12 and name H5" and segid 1) (resid 12 and name H1' and segid 1) 3.1 1.5 1.5
assign (resid 12 and name H2' and segid 1) (resid 12 and name H8 and segid 1) 4.5 1.8 1.8
assign (resid 12 and name H2" and segid 1) (resid 12 and name H8 and segid 1) 3.2 2.0 2.0
assign (resid 12 and name H2' and segid 1) (resid 12 and name H4' and segid 1) 3.7 1.5 1.5
assign (resid 12 and name H1' and segid 1) (resid 12 and name H8 and segid 1) 3.8 1.5 1.5
assign (resid 12 and name H2' and segid 1) (resid 12 and name H2" and segid 1) 2.7 1.1 1.1
assign (resid 12 and name H4' and segid 1) (resid 12 and name H1' and segid 1) 4.0 1.6 1.6
assign (resid 14 and name H1' and segid 1) (resid 14 and name H6 and segid 1) 2.2 1.5 1.5

assign (resid 14 and name H2' and segid 1) (resid 14 and name H5" and segid 1) 3.1 1.2 1.2
assign (resid 14 and name H2" and segid 1) (resid 14 and name H1' and segid 1) 3.3 1.3 1.3
assign (resid 14 and name H2" and segid 1) (resid 14 and name H4' and segid 1) 2.7 2.0 2.0
assign (resid 14 and name H5 and segid 1) (resid 14 and name H6 and segid 1) 2.7 1.1 1.1
assign (resid 14 and name H2' and segid 1) (resid 14 and name H2" and segid 1) 1.4 0.6 0.6
assign (resid 14 and name H2' and segid 1) (resid 14 and name H6 and segid 1) 1.7 1.7 1.7
assign (resid 14 and name H4' and segid 1) (resid 14 and name H6 and segid 1) 3.7 1.5 1.5
assign (resid 14 and name H5" and segid 1) (resid 14 and name H6 and segid 1) 3.2 1.3 1.3
assign (resid 14 and name H2' and segid 1) (resid 14 and name H5' and segid 1) 3.7 1.5 1.5
assign (resid 14 and name H2" and segid 1) (resid 14 and name H5' and segid 1) 3.7 1.5 1.5
assign (resid 14 and name H2' and segid 1) (resid 14 and name H1' and segid 1) 4.0 1.6 1.6
assign (resid 14 and name H5" and segid 1) (resid 15 and name H6 and segid 1) 3.6 1.4 1.4
assign (resid 15 and name H1' and segid 1) (resid 15 and name H6 and segid 1) 3.2 1.3 1.3
assign (resid 15 and name H2' and segid 1) (resid 15 and name H1' and segid 1) 3.5 1.4 1.4
assign (resid 15 and name H2' and segid 1) (resid 15 and name H2" and segid 1) 2.0 0.8 0.8
assign (resid 15 and name H2" and segid 1) (resid 15 and name H1' and segid 1) 1.6 0.6 0.6
assign (resid 15 and name H4' and segid 1) (resid 15 and name H6 and segid 1) 3.4 2.5 2.5
assign (resid 15 and name H5 and segid 1) (resid 15 and name H6 and segid 1) 2.2 0.9 0.9
assign (resid 15 and name H5" and segid 1) (resid 15 and name H6 and segid 1) 3.0 2.5 2.5
assign (resid 15 and name H5' and segid 1) (resid 15 and name H6 and segid 1) 2.7 2.0 2.0
assign (resid 15 and name H5" and segid 1) (resid 15 and name H1' and segid 1) 4.0 1.6 1.6
assign (resid 15 and name H2' and segid 1) (resid 15 and name H6 and segid 1) 2.6 2.0 2.0
assign (resid 15 and name H2" and segid 1) (resid 16 and name H6 and segid 1) 2.4 1.5 1.5
assign (resid 16 and name H7* and segid 1) (resid 15 and name H6 and segid 1) 2.6 1.5 1.5
assign (resid 16 and name H7* and segid 1) (resid 15 and name H5 and segid 1) 2.9 2.5 2.5
assign (resid 16 and name H1' and segid 1) (resid 16 and name H6 and segid 1) 3.5 1.4 1.4
assign (resid 16 and name H2' and segid 1) (resid 16 and name H2" and segid 1) 2.5 1.0 1.0
assign (resid 16 and name H2" and segid 1) (resid 16 and name H4' and segid 1) 3.2 1.3 1.3
assign (resid 16 and name H2" and segid 1) (resid 17 and name H6 and segid 1) 2.0 2.0 2.0
assign (resid 16 and name H7* and segid 1) (resid 16 and name H6 and segid 1) 2.0 0.8 0.8
assign (resid 16 and name H2' and segid 1) (resid 16 and name H5" and segid 1) 3.1 1.2 1.2
assign (resid 16 and name H4' and segid 1) (resid 16 and name H1' and segid 1) 2.8 1.1 1.1
assign (resid 16 and name H2' and segid 1) (resid 16 and name H6 and segid 1) 1.4 1.4 1.4
assign (resid 16 and name H2" and segid 1) (resid 16 and name H1' and segid 1) 3.1 1.2 1.2
assign (resid 16 and name H2" and segid 1) (resid 16 and name H6 and segid 1) 3.0 1.2 1.2
assign (resid 17 and name H7* and segid 1) (resid 17 and name H6 and segid 1) 2.7 1.1 1.1
assign (resid 17 and name H7* and segid 1) (resid 17 and name H1' and segid 1) 3.1 2.0 2.0
assign (resid 17 and name H2' and segid 1) (resid 17 and name H1' and segid 1) 2.8 1.1 1.1
assign (resid 17 and name H2' and segid 1) (resid 17 and name H2" and segid 1) 1.4 0.5 0.5
assign (resid 17 and name H1' and segid 1) (resid 17 and name H6 and segid 1) 3.1 1.2 1.2
assign (resid 17 and name H2" and segid 1) (resid 17 and name H1' and segid 1) 2.6 1.0 1.0
assign (resid 17 and name H7* and segid 1) (resid 18 and name H5 and segid 1) 3.4 2.5 3.0
assign (resid 18 and name H2' and segid 1) (resid 18 and name H2" and segid 1) 2.8 1.1 1.1
assign (resid 18 and name H1' and segid 1) (resid 18 and name H6 and segid 1) 3.0 1.2 1.2
assign (resid 18 and name H2' and segid 1) (resid 18 and name H6 and segid 1) 1.8 1.5 2.0
assign (resid 18 and name H5 and segid 1) (resid 18 and name H6 and segid 1) 2.4 1.0 1.0
assign (resid 19 and name H2" and segid 1) (resid 19 and name H6 and segid 1) 2.4 1.5 1.5
assign (resid 19 and name H5 and segid 1) (resid 19 and name H6 and segid 1) 2.3 1.5 1.5
assign (resid 19 and name H2' and segid 1) (resid 19 and name H1' and segid 1) 2.8 1.1 1.1
assign (resid 19 and name H2" and segid 1) (resid 19 and name H1' and segid 1) 2.7 1.1 1.1
assign (resid 19 and name H2' and segid 1) (resid 19 and name H6 and segid 1) 2.1 1.5 2.0
assign (resid 20 and name H7* and segid 1) (resid 19 and name H6 and segid 1) 2.6 1.0 2.0
assign (resid 20 and name H7* and segid 1) (resid 19 and name H5 and segid 1) 2.9 2.0 2.0
assign (resid 20 and name H5' and segid 1) (resid 20 and name H6 and segid 1) 3.1 1.2 1.2
assign (resid 20 and name H1' and segid 1) (resid 20 and name H6 and segid 1) 2.8 1.1 1.1
assign (resid 20 and name H2' and segid 1) (resid 20 and name H2" and segid 1) 2.1 0.9 0.9
assign (resid 20 and name H2" and segid 1) (resid 20 and name H1' and segid 1) 1.8 0.7 0.7
assign (resid 20 and name H2' and segid 1) (resid 20 and name H6 and segid 1) 2.5 1.5 1.5
assign (resid 20 and name H7* and segid 1) (resid 20 and name H6 and segid 1) 2.1 0.8 0.8
assign (resid 21 and name H2' and segid 1) (resid 21 and name H2" and segid 1) 3.0 1.2 1.2

assign (resid 21 and name H1' and segid 1) (resid 21 and name H6 and segid 1) 3.3 1.3 1.3
assign (resid 21 and name H2' and segid 1) (resid 21 and name H5' and segid 1) 2.7 1.1 1.1
assign (resid 21 and name H2' and segid 1) (resid 21 and name H6 and segid 1) 2.2 0.9 0.9
assign (resid 21 and name H2" and segid 1) (resid 21 and name H1' and segid 1) 2.5 1.0 1.0
assign (resid 21 and name H7* and segid 1) (resid 21 and name H1' and segid 1) 3.2 2.0 2.0
assign (resid 21 and name H7* and segid 1) (resid 21 and name H6 and segid 1) 2.3 0.9 0.9
assign (resid 21 and name H2' and segid 1) (resid 21 and name H1' and segid 1) 2.1 0.8 0.8
assign (resid 21 and name H2" and segid 1) (resid 21 and name H6 and segid 1) 3.0 1.2 1.2
assign (resid 22 and name H5 and segid 1) (resid 21 and name H6 and segid 1) 3.2 2.5 2.5
assign (resid 21 and name H7* and segid 1) (resid 22 and name H5 and segid 1) 3.0 2.0 2.0
assign (resid 22 and name H1' and segid 1) (resid 22 and name H6 and segid 1) 2.9 1.2 1.2
assign (resid 22 and name H4' and segid 1) (resid 22 and name H6 and segid 1) 3.5 2.0 2.0
assign (resid 22 and name H2' and segid 1) (resid 22 and name H6 and segid 1) 1.3 1.3 2.0
assign (resid 22 and name H5 and segid 1) (resid 22 and name H6 and segid 1) 2.3 2.0 2.0
assign (resid 22 and name H2" and segid 1) (resid 22 and name H6 and segid 1) 2.2 2.0 2.0
assign (resid 23 and name H2' and segid 1) (resid 23 and name H1' and segid 1) 3.9 1.6 1.6
assign (resid 23 and name H2' and segid 1) (resid 23 and name H2" and segid 1) 2.1 0.8 0.8
assign (resid 23 and name H2' and segid 1) (resid 23 and name H6 and segid 1) 2.8 1.1 1.1
assign (resid 23 and name H2" and segid 1) (resid 23 and name H1' and segid 1) 2.3 0.9 0.9
assign (resid 23 and name H4' and segid 1) (resid 23 and name H1' and segid 1) 1.8 1.0 1.0
assign (resid 23 and name H5 and segid 1) (resid 23 and name H6 and segid 1) 2.4 1.0 1.0
assign (resid 23 and name H1' and segid 1) (resid 23 and name H6 and segid 1) 2.4 1.6 1.6
assign (resid 23 and name H2" and segid 1) (resid 23 and name H6 and segid 1) 2.2 1.5 1.5
assign (resid 23 and name H5' and segid 1) (resid 23 and name H1' and segid 1) 4.3 1.7 1.7
assign (resid 24 and name H4' and segid 1) (resid 24 and name H6 and segid 1) 3.6 1.4 1.4
assign (resid 24 and name H1' and segid 1) (resid 24 and name H6 and segid 1) 2.9 1.2 1.2
assign (resid 24 and name H4' and segid 1) (resid 24 and name H1' and segid 1) 2.4 2.0 2.0
assign (resid 24 and name H7* and segid 1) (resid 23 and name H1' and segid 1) 2.9 1.2 1.2
assign (resid 24 and name H7* and segid 1) (resid 23 and name H5 and segid 1) 2.8 2.0 2.0
assign (resid 24 and name H7* and segid 1) (resid 24 and name H6 and segid 1) 2.1 0.8 0.8
assign (resid 24 and name H2" and segid 1) (resid 24 and name H6 and segid 1) 2.9 2.0 2.5
assign (resid 24 and name H3' and segid 1) (resid 24 and name H1' and segid 1) 3.5 1.4 1.4
assign (resid 24 and name H5' and segid 1) (resid 24 and name H6 and segid 1) 2.7 1.1 1.1
assign (resid 24 and name H2' and segid 1) (resid 25 and name H6 and segid 1) 3.0 2.0 2.0
assign (resid 25 and name H2" and segid 1) (resid 25 and name H6 and segid 1) 2.5 2.0 2.0
assign (resid 25 and name H4' and segid 1) (resid 25 and name H6 and segid 1) 4.1 2.5 2.5
assign (resid 25 and name H2' and segid 1) (resid 25 and name H1' and segid 1) 2.1 1.2 1.2
assign (resid 25 and name H3' and segid 1) (resid 25 and name H1' and segid 1) 3.5 2.1 2.1
assign (resid 25 and name H4' and segid 1) (resid 25 and name H3' and segid 1) 2.9 1.7 1.7
assign (resid 25 and name H5' and segid 1) (resid 25 and name H6 and segid 1) 2.7 1.6 1.6
assign (resid 25 and name H1' and segid 1) (resid 25 and name H6 and segid 1) 3.0 1.8 1.8
assign (resid 25 and name H2' and segid 1) (resid 25 and name H3' and segid 1) 2.4 1.4 1.4
assign (resid 25 and name H2' and segid 1) (resid 25 and name H5' and segid 1) 2.8 2.5 2.5
assign (resid 25 and name H2' and segid 1) (resid 25 and name H6 and segid 1) 1.7 1.0 1.0
assign (resid 25 and name H3' and segid 1) (resid 25 and name H6 and segid 1) 2.9 1.7 1.7
assign (resid 25 and name H5' and segid 1) (resid 25 and name H1' and segid 1) 2.8 1.7 1.7
assign (resid 25 and name H4' and segid 1) (resid 25 and name H1' and segid 1) 3.7 2.2 2.2

Exchangeable proton distance restraints file

assign (resid 4 and name H1 and segid 1) (resid 3 and name H1 and segid 1) 3.0 2.5 3.0
assign (resid 4 and name H1 and segid 1) (resid 21 and name H3 and segid 1) 3.0 1.5 2.5
assign (resid 5 and name H2 and segid 1) (resid 21 and name H3 and segid 1) 2.2 2.2 2.2
assign (resid 6 and name H2 and segid 1) (resid 20 and name H3 and segid 1) 3.0 2.5 2.5
assign (resid 7 and name H2 and segid 1) (resid 19 and name H3 and segid 1) 3.1 2.5 2.5
assign (resid 8 and name H1 and segid 1) (resid 17 and name H3 and segid 1) 3.1 2.5 3.0
assign (resid 9 and name H2 and segid 1) (resid 17 and name H3 and segid 1) 3.3 2.0 3.0

assign (resid 16 and name H3 and segid 1) (resid 17 and name H3 and segid 1) 3.0 2.5 2.5
assign (resid 17 and name H7* and segid 1) (resid 17 and name H3 and segid 1) 3.1 1.6 1.6
assign (resid 18 and name H5 and segid 1) (resid 8 and name H1 and segid 1) 2.9 2.0 2.5
assign (resid 18 and name H41 and segid 1) (resid 8 and name H1 and segid 1) 2.2 1.1 1.1
assign (resid 19 and name H5 and segid 1) (resid 19 and name H3 and segid 1) 3.1 1.6 1.6
assign (resid 20 and name H7* and segid 1) (resid 19 and name H3 and segid 1) 3.1 2.0 2.0
assign (resid 20 and name H7* and segid 1) (resid 20 and name H3 and segid 1) 3.0 1.5 1.5
assign (resid 21 and name H3 and segid 1) (resid 20 and name H3 and segid 1) 2.7 2.5 2.5
assign (resid 21 and name H7* and segid 1) (resid 4 and name H1 and segid 1) 3.1 2.5 2.5
assign (resid 21 and name H7* and segid 1) (resid 21 and name H3 and segid 1) 3.0 1.5 1.5
assign (resid 22 and name H5 and segid 1) (resid 3 and name H1 and segid 1) 3.1 2.5 2.5
assign (resid 22 and name H5 and segid 1) (resid 4 and name H1 and segid 1) 2.9 1.5 1.5
assign (resid 22 and name H41 and segid 1) (resid 4 and name H1 and segid 1) 2.3 1.2 1.2
assign (resid 23 and name H5 and segid 1) (resid 3 and name H1 and segid 1) 3.0 1.5 1.5
assign (resid 23 and name H41 and segid 1) (resid 3 and name H1 and segid 1) 2.5 1.3 1.3

A2. NOe distance restraints files for non-modified triplex

Non-exchangeable proton distance restraints file

assign (resid 1 and name H2' and segid 1) (resid 1 and name H1' and segid 1) 2.2 1.1 1.1
assign (resid 1 and name H2' and segid 1) (resid 1 and name H8 and segid 1) 2.9 1.8 1.8
assign (resid 1 and name H2" and segid 1) (resid 1 and name H4' and segid 1) 2.5 1.5 1.5
assign (resid 1 and name H2" and segid 1) (resid 1 and name H1' and segid 1) 2.5 1.3 1.3
assign (resid 1 and name H1' and segid 1) (resid 1 and name H8 and segid 1) 3.1 1.6 1.6
assign (resid 1 and name H2' and segid 1) (resid 1 and name H2" and segid 1) 2.0 1.0 1.0
assign (resid 1 and name H2" and segid 1) (resid 1 and name H8 and segid 1) 3.2 1.6 1.6
assign (resid 1 and name H4' and segid 1) (resid 1 and name H1' and segid 1) 3.7 1.9 1.9
assign (resid 1 and name H4' and segid 1) (resid 1 and name H8 and segid 1) 3.2 1.6 1.6
assign (resid 1 and name H5' and segid 1) (resid 1 and name H8 and segid 1) 3.6 1.8 1.8
assign (resid 1 and name H5" and segid 1) (resid 1 and name H4' and segid 1) 1.7 0.9 0.9
assign (resid 2 and name H8 and segid 1) (resid 1 and name H8 and segid 1) 4.4 2.2 2.2
assign (resid 2 and name H2' and segid 1) (resid 1 and name H8 and segid 1) 3.3 1.6 1.6
assign (resid 2 and name H1' and segid 1) (resid 1 and name H8 and segid 1) 3.1 2.3 2.3
assign (resid 2 and name H1' and segid 1) (resid 2 and name H8 and segid 1) 3.0 1.5 1.5
assign (resid 2 and name H2' and segid 1) (resid 2 and name H1' and segid 1) 2.4 1.0 1.0
assign (resid 2 and name H2' and segid 1) (resid 2 and name H2" and segid 1) 2.2 0.9 0.9
assign (resid 2 and name H2' and segid 1) (resid 2 and name H8 and segid 1) 2.4 2.0 2.0
assign (resid 2 and name H2" and segid 1) (resid 2 and name H1' and segid 1) 2.7 1.1 1.1
assign (resid 2 and name H4' and segid 1) (resid 2 and name H1' and segid 1) 3.1 1.2 1.2
assign (resid 2 and name H5' and segid 1) (resid 2 and name H4' and segid 1) 1.9 0.8 0.8
assign (resid 2 and name H5' and segid 1) (resid 2 and name H8 and segid 1) 3.3 3.0 3.0
assign (resid 2 and name H2' and segid 1) (resid 2 and name H4' and segid 1) 4.6 1.8 1.8
assign (resid 2 and name H2" and segid 1) (resid 2 and name H8 and segid 1) 3.2 1.3 1.3
assign (resid 2 and name H4' and segid 1) (resid 2 and name H8 and segid 1) 4.8 1.9 1.9
assign (resid 3 and name H1' and segid 1) (resid 2 and name H8 and segid 1) 3.8 2.0 2.0
assign (resid 3 and name H2' and segid 1) (resid 2 and name H8 and segid 1) 2.9 2.0 2.0
assign (resid 2 and name H8 and segid 1) (resid 3 and name H8 and segid 1) 3.9 2.5 2.5
assign (resid 3 and name H2' and segid 1) (resid 3 and name H1' and segid 1) 2.9 1.2 1.2
assign (resid 3 and name H1' and segid 1) (resid 3 and name H8 and segid 1) 3.1 1.2 1.2
assign (resid 3 and name H2' and segid 1) (resid 3 and name H2" and segid 1) 2.5 1.0 1.0

assign (resid 3 and name H2' and segid 1) (resid 3 and name H8 and segid 1) 2.4 1.5 1.5
assign (resid 3 and name H2" and segid 1) (resid 3 and name H1' and segid 1) 2.9 1.2 1.2
assign (resid 3 and name H2' and segid 1) (resid 3 and name H4' and segid 1) 3.4 1.4 1.4
assign (resid 3 and name H2" and segid 1) (resid 3 and name H8 and segid 1) 3.5 1.4 1.4
assign (resid 3 and name H3' and segid 1) (resid 3 and name H1' and segid 1) 3.7 1.5 1.5
assign (resid 3 and name H4' and segid 1) (resid 3 and name H8 and segid 1) 3.3 2.5 2.5
assign (resid 3 and name H4' and segid 1) (resid 3 and name H1' and segid 1) 2.7 1.1 1.1
assign (resid 3 and name H8 and segid 1) (resid 4 and name H8 and segid 1) 4.0 3.0 3.0
assign (resid 4 and name H2' and segid 1) (resid 4 and name H2" and segid 1) 2.2 0.9 0.9
assign (resid 4 and name H2' and segid 1) (resid 4 and name H3' and segid 1) 2.5 1.0 1.0
assign (resid 4 and name H2' and segid 1) (resid 4 and name H4' and segid 1) 3.3 1.3 1.3
assign (resid 4 and name H2' and segid 1) (resid 4 and name H8 and segid 1) 2.6 1.5 1.5
assign (resid 4 and name H2" and segid 1) (resid 4 and name H1' and segid 1) 2.3 0.9 0.9
assign (resid 4 and name H2" and segid 1) (resid 4 and name H3' and segid 1) 2.2 0.9 0.9
assign (resid 4 and name H2" and segid 1) (resid 4 and name H4' and segid 1) 2.7 1.5 1.5
assign (resid 4 and name H5' and segid 1) (resid 4 and name H4' and segid 1) 2.0 0.8 0.8
assign (resid 4 and name H1' and segid 1) (resid 4 and name H8 and segid 1) 3.8 1.5 1.5
assign (resid 4 and name H3' and segid 1) (resid 4 and name H1' and segid 1) 4.0 1.6 1.6
assign (resid 4 and name H5" and segid 1) (resid 4 and name H4' and segid 1) 2.0 0.8 0.8
assign (resid 4 and name H4' and segid 1) (resid 4 and name H8 and segid 1) 3.7 2.0 2.0
assign (resid 4 and name H2' and segid 1) (resid 4 and name H1' and segid 1) 2.5 1.0 1.0
assign (resid 4 and name H2" and segid 1) (resid 4 and name H8 and segid 1) 2.8 1.5 1.5
assign (resid 4 and name H4' and segid 1) (resid 4 and name H1' and segid 1) 3.1 1.2 1.2
assign (resid 4 and name H5" and segid 1) (resid 4 and name H8 and segid 1) 3.7 3.0 3.0
assign (resid 4 and name H5' and segid 1) (resid 4 and name H8 and segid 1) 3.9 2.5 2.5
assign (resid 4 and name H3' and segid 1) (resid 4 and name H8 and segid 1) 3.5 2.0 2.0
assign (resid 5 and name H8 and segid 1) (resid 4 and name H8 and segid 1) 4.5 1.8 1.8
assign (resid 5 and name H1' and segid 1) (resid 4 and name H8 and segid 1) 3.4 2.0 2.0
assign (resid 5 and name H2' and segid 1) (resid 4 and name H8 and segid 1) 3.2 1.3 1.3
assign (resid 5 and name H2" and segid 1) (resid 4 and name H8 and segid 1) 3.2 2.0 2.0
assign (resid 5 and name H2' and segid 1) (resid 5 and name H1' and segid 1) 2.9 1.2 1.2
assign (resid 5 and name H2" and segid 1) (resid 5 and name H1' and segid 1) 2.6 1.0 1.0
assign (resid 5 and name H4' and segid 1) (resid 5 and name H1' and segid 1) 3.1 1.2 1.2
assign (resid 5 and name H5" and segid 1) (resid 5 and name H1' and segid 1) 3.3 2.0 2.0
assign (resid 5 and name H2' and segid 1) (resid 5 and name H4' and segid 1) 2.8 1.1 1.1
assign (resid 5 and name H2" and segid 1) (resid 5 and name H4' and segid 1) 2.3 1.5 1.5
assign (resid 5 and name H5' and segid 1) (resid 5 and name H1' and segid 1) 3.3 1.3 1.3
assign (resid 5 and name H1' and segid 1) (resid 5 and name H8 and segid 1) 1.9 1.2 1.2
assign (resid 5 and name H2' and segid 1) (resid 5 and name H2" and segid 1) 2.4 1.0 1.0
assign (resid 5 and name H2' and segid 1) (resid 5 and name H8 and segid 1) 3.2 1.3 1.3
assign (resid 5 and name H3' and segid 1) (resid 5 and name H1' and segid 1) 2.9 1.2 1.2
assign (resid 5 and name H4' and segid 1) (resid 5 and name H8 and segid 1) 3.5 2.0 2.0
assign (resid 5 and name H8 and segid 1) (resid 6 and name H8 and segid 1) 3.8 1.5 1.5
assign (resid 6 and name H2" and segid 1) (resid 5 and name H8 and segid 1) 3.0 2.0 2.0
assign (resid 6 and name H1' and segid 1) (resid 5 and name H8 and segid 1) 5.6 2.2 2.2
assign (resid 6 and name H2' and segid 1) (resid 5 and name H8 and segid 1) 3.0 1.5 1.5
assign (resid 6 and name H3' and segid 1) (resid 5 and name H8 and segid 1) 3.7 1.5 1.5
assign (resid 6 and name H2' and segid 1) (resid 6 and name H2" and segid 1) 2.7 1.1 1.1
assign (resid 6 and name H2' and segid 1) (resid 6 and name H3' and segid 1) 3.3 1.3 1.3
assign (resid 6 and name H2' and segid 1) (resid 6 and name H8 and segid 1) 3.0 1.2 1.2
assign (resid 6 and name H2" and segid 1) (resid 6 and name H1' and segid 1) 3.3 1.3 1.3
assign (resid 6 and name H2" and segid 1) (resid 6 and name H8 and segid 1) 2.5 1.0 1.0
assign (resid 6 and name H4' and segid 1) (resid 6 and name H3' and segid 1) 2.7 1.1 1.1
assign (resid 6 and name H2' and segid 1) (resid 6 and name H4' and segid 1) 3.5 1.4 1.4
assign (resid 6 and name H3' and segid 1) (resid 6 and name H8 and segid 1) 3.1 2.0 2.0
assign (resid 6 and name H4' and segid 1) (resid 6 and name H1' and segid 1) 2.8 1.1 1.1
assign (resid 6 and name H1' and segid 1) (resid 6 and name H8 and segid 1) 2.6 1.0 1.0
assign (resid 6 and name H3' and segid 1) (resid 6 and name H1' and segid 1) 3.4 1.4 1.4
assign (resid 6 and name H4' and segid 1) (resid 6 and name H8 and segid 1) 3.7 2.0 2.0
assign (resid 6 and name H5' and segid 1) (resid 6 and name H1' and segid 1) 3.5 1.4 1.4

assign (resid 6 and name H8 and segid 1) (resid 7 and name H8 and segid 1) 4.0 2.0 2.0
assign (resid 7 and name H1' and segid 1) (resid 6 and name H8 and segid 1) 3.7 3.0 3.0
assign (resid 7 and name H2' and segid 1) (resid 7 and name H2" and segid 1) 2.1 0.8 0.8
assign (resid 7 and name H2" and segid 1) (resid 7 and name H1' and segid 1) 2.0 0.8 0.8
assign (resid 7 and name H2" and segid 1) (resid 7 and name H8 and segid 1) 2.6 1.5 1.5
assign (resid 7 and name H4' and segid 1) (resid 7 and name H1' and segid 1) 4 1.6 1.6
assign (resid 7 and name H4' and segid 1) (resid 7 and name H3' and segid 1) 2.5 1.0 1.0
assign (resid 7 and name H5' and segid 1) (resid 7 and name H1' and segid 1) 4.2 1.7 1.7
assign (resid 7 and name H5' and segid 1) (resid 7 and name H3' and segid 1) 2.7 1.1 1.1
assign (resid 7 and name H5' and segid 1) (resid 7 and name H4' and segid 1) 2.4 1.0 1.0
assign (resid 7 and name H3' and segid 1) (resid 7 and name H8 and segid 1) 3.7 2.0 2.0
assign (resid 7 and name H5" and segid 1) (resid 7 and name H3' and segid 1) 2.9 1.2 1.2
assign (resid 7 and name H3' and segid 1) (resid 7 and name H1' and segid 1) 3.1 1.2 1.2
assign (resid 7 and name H5' and segid 1) (resid 7 and name H8 and segid 1) 3.7 2.5 2.5
assign (resid 7 and name H2' and segid 1) (resid 7 and name H1' and segid 1) 2.7 1.1 1.1
assign (resid 7 and name H2' and segid 1) (resid 7 and name H8 and segid 1) 2.6 1.5 1.5
assign (resid 7 and name H4' and segid 1) (resid 7 and name H8 and segid 1) 4.6 1.8 1.8
assign (resid 7 and name H3' and segid 1) (resid 7 and name H1' and segid 1) 4.7 1.9 1.9
assign (resid 8 and name H2' and segid 1) (resid 8 and name H1' and segid 1) 2.7 1.5 1.5
assign (resid 8 and name H2' and segid 1) (resid 8 and name H2" and segid 1) 2.9 1.2 1.2
assign (resid 8 and name H2' and segid 1) (resid 8 and name H8 and segid 1) 2.5 1.5 1.5
assign (resid 8 and name H2" and segid 1) (resid 8 and name H8 and segid 1) 3.1 1.2 1.2
assign (resid 8 and name H1' and segid 1) (resid 8 and name H8 and segid 1) 3.3 1.3 1.3
assign (resid 8 and name H2" and segid 1) (resid 8 and name H1' and segid 1) 1.4 1.0 1.0
assign (resid 8 and name H3' and segid 1) (resid 8 and name H1' and segid 1) 3.1 1.2 1.2
assign (resid 8 and name H3' and segid 1) (resid 8 and name H8 and segid 1) 3.5 2.0 2.0
assign (resid 8 and name H4' and segid 1) (resid 8 and name H8 and segid 1) 3.6 2.0 2.0
assign (resid 9 and name H8 and segid 1) (resid 8 and name H8 and segid 1) 3.3 1.3 1.3
assign (resid 9 and name H2" and segid 1) (resid 8 and name H8 and segid 1) 3.2 2.5 2.5
assign (resid 9 and name H1' and segid 1) (resid 8 and name H8 and segid 1) 3.2 2.5 2.5
assign (resid 9 and name H2' and segid 1) (resid 8 and name H8 and segid 1) 3.2 2.0 2.0
assign (resid 9 and name H2' and segid 1) (resid 9 and name H1' and segid 1) 3.3 1.3 1.3
assign (resid 9 and name H2' and segid 1) (resid 9 and name H2" and segid 1) 2.0 0.8 0.8
assign (resid 9 and name H2" and segid 1) (resid 9 and name H1' and segid 1) 2.7 1.1 1.1
assign (resid 9 and name H4' and segid 1) (resid 9 and name H1' and segid 1) 3.2 1.3 1.3
assign (resid 9 and name H4' and segid 1) (resid 9 and name H3' and segid 1) 2.5 1.0 1.0
assign (resid 9 and name H5" and segid 1) (resid 9 and name H1' and segid 1) 3.1 2.0 2.0
assign (resid 9 and name H2" and segid 1) (resid 9 and name H8 and segid 1) 2.7 1.5 1.5
assign (resid 9 and name H3' and segid 1) (resid 9 and name H1' and segid 1) 3.5 1.4 1.4
assign (resid 9 and name H5' and segid 1) (resid 9 and name H1' and segid 1) 2.9 2.0 2.0
assign (resid 9 and name H1' and segid 1) (resid 9 and name H8 and segid 1) 3.1 1.2 1.2
assign (resid 9 and name H3' and segid 1) (resid 9 and name H8 and segid 1) 4.2 1.7 1.7
assign (resid 9 and name H4' and segid 1) (resid 9 and name H8 and segid 1) 4.2 1.7 1.7
assign (resid 10 and name H2" and segid 1) (resid 10 and name H8 and segid 1) 2.9 1.2 1.2
assign (resid 10 and name H1' and segid 1) (resid 10 and name H8 and segid 1) 3.4 1.4 1.4
assign (resid 10 and name H2' and segid 1) (resid 10 and name H3' and segid 1) 2.2 0.9 0.9
assign (resid 10 and name H2" and segid 1) (resid 10 and name H1' and segid 1) 2.4 1.0 1.0
assign (resid 10 and name H2" and segid 1) (resid 10 and name H3' and segid 1) 2.1 0.8 0.8
assign (resid 10 and name H2" and segid 1) (resid 10 and name H4' and segid 1) 2.6 1.5 1.5
assign (resid 10 and name H4' and segid 1) (resid 10 and name H1' and segid 1) 3.2 1.3 1.3
assign (resid 10 and name H2' and segid 1) (resid 10 and name H1' and segid 1) 2.4 1.0 1.0
assign (resid 10 and name H2' and segid 1) (resid 10 and name H4' and segid 1) 2.7 1.1 1.1
assign (resid 10 and name H3' and segid 1) (resid 10 and name H8 and segid 1) 3.9 1.6 1.6
assign (resid 10 and name H3' and segid 1) (resid 10 and name H1' and segid 1) 4.3 1.7 1.7
assign (resid 10 and name H4' and segid 1) (resid 10 and name H8 and segid 1) 4.0 1.6 1.6
assign (resid 11 and name H8 and segid 1) (resid 10 and name H8 and segid 1) 4.0 1.6 1.6
assign (resid 11 and name H3' and segid 1) (resid 10 and name H8 and segid 1) 4.6 1.8 1.8
assign (resid 11 and name H1' and segid 1) (resid 10 and name H8 and segid 1) 3.5 2.0 2.0
assign (resid 11 and name H2" and segid 1) (resid 10 and name H8 and segid 1) 3.1 2.5 2.5
assign (resid 11 and name H2' and segid 1) (resid 10 and name H8 and segid 1) 3.2 2.0 2.0

assign (resid 11 and name H2' and segid 1) (resid 11 and name H2" and segid 1) 2.3 0.9 0.9
assign (resid 11 and name H2' and segid 1) (resid 11 and name H3' and segid 1) 3.1 1.2 1.2
assign (resid 11 and name H2' and segid 1) (resid 11 and name H5' and segid 1) 2.3 1.5 1.5
assign (resid 11 and name H2' and segid 1) (resid 11 and name H8 and segid 1) 2.6 1.0 1.0
assign (resid 11 and name H2" and segid 1) (resid 11 and name H1' and segid 1) 2.7 1.1 1.1
assign (resid 11 and name H2" and segid 1) (resid 11 and name H3' and segid 1) 2.6 1.0 1.0
assign (resid 11 and name H2" and segid 1) (resid 11 and name H4' and segid 1) 4.1 1.6 1.6
assign (resid 11 and name H2" and segid 1) (resid 11 and name H8 and segid 1) 2.9 1.2 1.2
assign (resid 11 and name H5" and segid 1) (resid 11 and name H4' and segid 1) 2.9 1.2 1.2
assign (resid 11 and name H1' and segid 1) (resid 11 and name H8 and segid 1) 3.3 1.3 1.3
assign (resid 11 and name H2' and segid 1) (resid 11 and name H4' and segid 1) 2.7 1.1 1.1
assign (resid 11 and name H3' and segid 1) (resid 11 and name H8 and segid 1) 3.4 2.0 2.0
assign (resid 11 and name H2' and segid 1) (resid 11 and name H1' and segid 1) 4.0 1.6 1.6
assign (resid 11 and name H2" and segid 1) (resid 11 and name H5' and segid 1) 3.6 1.4 1.4
assign (resid 11 and name H3' and segid 1) (resid 11 and name H1' and segid 1) 3.4 1.4 1.4
assign (resid 11 and name H4' and segid 1) (resid 11 and name H8 and segid 1) 3.7 2.0 2.0
assign (resid 11 and name H5' and segid 1) (resid 11 and name H8 and segid 1) 3.4 2.5 2.5
assign (resid 12 and name H2' and segid 1) (resid 11 and name H8 and segid 1) 3.4 1.7 1.7
assign (resid 12 and name H2" and segid 1) (resid 11 and name H8 and segid 1) 3.2 1.6 1.6
assign (resid 12 and name H2' and segid 1) (resid 12 and name H5' and segid 1) 2.1 1.2 1.2
assign (resid 12 and name H2" and segid 1) (resid 12 and name H1' and segid 1) 2.8 1.4 1.4
assign (resid 12 and name H5' and segid 1) (resid 12 and name H1' and segid 1) 2.9 1.5 1.5
assign (resid 12 and name H5" and segid 1) (resid 12 and name H1' and segid 1) 3.2 2.0 2.0
assign (resid 12 and name H2' and segid 1) (resid 12 and name H8 and segid 1) 3.1 1.6 1.6
assign (resid 12 and name H2" and segid 1) (resid 12 and name H8 and segid 1) 3.5 1.8 1.8
assign (resid 12 and name H2' and segid 1) (resid 12 and name H4' and segid 1) 4.1 2.1 2.1
assign (resid 12 and name H1' and segid 1) (resid 12 and name H8 and segid 1) 3.8 1.9 1.9
assign (resid 12 and name H2' and segid 1) (resid 12 and name H2" and segid 1) 2.7 1.4 1.4
assign (resid 12 and name H4' and segid 1) (resid 12 and name H1' and segid 1) 4.0 2.0 2.0
assign (resid 14 and name H1' and segid 1) (resid 14 and name H6 and segid 1) 2.2 1.5 1.5
assign (resid 14 and name H2' and segid 1) (resid 14 and name H5" and segid 1) 3.1 1.6 1.6
assign (resid 14 and name H2" and segid 1) (resid 14 and name H1' and segid 1) 3.7 1.7 1.7
assign (resid 14 and name H2" and segid 1) (resid 14 and name H4' and segid 1) 2.6 1.3 1.3
assign (resid 14 and name H5 and segid 1) (resid 14 and name H6 and segid 1) 2.4 1.2 1.2
assign (resid 14 and name H2' and segid 1) (resid 14 and name H2" and segid 1) 1.6 0.8 0.8
assign (resid 14 and name H2' and segid 1) (resid 14 and name H6 and segid 1) 1.8 0.9 0.9
assign (resid 14 and name H4' and segid 1) (resid 14 and name H6 and segid 1) 3.3 2.0 2.0
assign (resid 14 and name H5" and segid 1) (resid 14 and name H6 and segid 1) 3.3 1.7 1.7
assign (resid 14 and name H2' and segid 1) (resid 14 and name H5' and segid 1) 3.3 1.7 1.7
assign (resid 14 and name H2" and segid 1) (resid 14 and name H5' and segid 1) 3.3 1.7 1.7
assign (resid 14 and name H2' and segid 1) (resid 14 and name H1' and segid 1) 4.0 2.0 2.0
assign (resid 14 and name H2" and segid 1) (resid 14 and name H5" and segid 1) 2.6 2.0 2.0
assign (resid 14 and name H5" and segid 1) (resid 15 and name H6 and segid 1) 3.6 1.7 1.7
assign (resid 15 and name H1' and segid 1) (resid 15 and name H6 and segid 1) 2.6 1.0 1.0
assign (resid 15 and name H2' and segid 1) (resid 15 and name H1' and segid 1) 2.6 1.0 1.0
assign (resid 15 and name H2' and segid 1) (resid 15 and name H2" and segid 1) 2.1 0.8 0.8
assign (resid 15 and name H2" and segid 1) (resid 15 and name H1' and segid 1) 1.4 1.0 1.0
assign (resid 15 and name H4' and segid 1) (resid 15 and name H6 and segid 1) 3.6 1.4 1.4
assign (resid 15 and name H5 and segid 1) (resid 15 and name H6 and segid 1) 2.4 1.0 1.0
assign (resid 15 and name H5" and segid 1) (resid 15 and name H6 and segid 1) 3.9 1.6 1.6
assign (resid 15 and name H5' and segid 1) (resid 15 and name H6 and segid 1) 3.5 1.4 1.4
assign (resid 15 and name H2' and segid 1) (resid 15 and name H6 and segid 1) 2.6 1.5 1.5
assign (resid 15 and name H2" and segid 1) (resid 15 and name H6 and segid 1) 2.7 2.0 2.0
assign (resid 15 and name H2" and segid 1) (resid 16 and name H6 and segid 1) 2.4 1.0 1.0
assign (resid 16 and name H7* and segid 1) (resid 15 and name H1' and segid 1) 2.8 1.1 1.1
assign (resid 16 and name H7* and segid 1) (resid 15 and name H6 and segid 1) 2.8 1.1 1.1
assign (resid 16 and name H7* and segid 1) (resid 15 and name H5 and segid 1) 3.0 2.0 2.0
assign (resid 16 and name H1' and segid 1) (resid 16 and name H6 and segid 1) 3.8 1.5 1.5
assign (resid 16 and name H2' and segid 1) (resid 16 and name H2" and segid 1) 3.0 1.2 1.2
assign (resid 16 and name H2' and segid 1) (resid 16 and name H6 and segid 1) 1.4 1.0 1.0

assign (resid 16 and name H2" and segid 1) (resid 16 and name H4' and segid 1) 3.2 1.3 1.3
assign (resid 16 and name H2" and segid 1) (resid 16 and name H6 and segid 1) 2.9 1.2 1.2
assign (resid 16 and name H7* and segid 1) (resid 16 and name H1' and segid 1) 3.1 2.0 2.0
assign (resid 16 and name H7* and segid 1) (resid 16 and name H6 and segid 1) 2.1 0.8 0.8
assign (resid 16 and name H2' and segid 1) (resid 16 and name H1' and segid 1) 1.9 1.0 1.0
assign (resid 16 and name H2' and segid 1) (resid 16 and name H5" and segid 1) 3.1 2.0 2.0
assign (resid 16 and name H4' and segid 1) (resid 16 and name H1' and segid 1) 2.2 1.5 1.5
assign (resid 16 and name H2" and segid 1) (resid 16 and name H1' and segid 1) 3.1 1.2 1.2
assign (resid 16 and name H2" and segid 1) (resid 17 and name H6 and segid 1) 3.8 1.5 1.5
assign (resid 17 and name H7* and segid 1) (resid 17 and name H6 and segid 1) 2.7 1.1 1.1
assign (resid 17 and name H2' and segid 1) (resid 17 and name H1' and segid 1) 2.8 1.1 1.1
assign (resid 17 and name H2' and segid 1) (resid 17 and name H2" and segid 1) 1.7 0.7 0.7
assign (resid 17 and name H1' and segid 1) (resid 17 and name H6 and segid 1) 2.1 1.5 1.5
assign (resid 17 and name H2" and segid 1) (resid 17 and name H1' and segid 1) 2.6 1.0 1.0
assign (resid 17 and name H7* and segid 1) (resid 18 and name H5 and segid 1) 3.4 1.4 1.4
assign (resid 18 and name H2' and segid 1) (resid 18 and name H2" and segid 1) 2.8 1.1 1.1
assign (resid 18 and name H1' and segid 1) (resid 18 and name H6 and segid 1) 3.9 1.6 1.6
assign (resid 18 and name H2' and segid 1) (resid 18 and name H6 and segid 1) 2.3 1.5 1.5
assign (resid 18 and name H2" and segid 1) (resid 18 and name H6 and segid 1) 4.0 1.6 1.6
assign (resid 18 and name H5 and segid 1) (resid 18 and name H6 and segid 1) 2.5 1.5 1.5
assign (resid 19 and name H2" and segid 1) (resid 19 and name H6 and segid 1) 2.4 1.5 1.5
assign (resid 19 and name H2" and segid 1) (resid 19 and name H1' and segid 1) 2.7 1.1 1.1
assign (resid 19 and name H5 and segid 1) (resid 19 and name H6 and segid 1) 2.6 1.0 1.0
assign (resid 19 and name H2' and segid 1) (resid 19 and name H1' and segid 1) 2.7 1.1 1.1
assign (resid 19 and name H2' and segid 1) (resid 19 and name H6 and segid 1) 2.1 2.0 2.0
assign (resid 20 and name H7* and segid 1) (resid 19 and name H6 and segid 1) 3.8 1.8 1.8
assign (resid 20 and name H7* and segid 1) (resid 19 and name H5 and segid 1) 3.0 2.0 2.0
assign (resid 20 and name H1' and segid 1) (resid 20 and name H6 and segid 1) 2.8 1.1 1.1
assign (resid 20 and name H2' and segid 1) (resid 20 and name H1' and segid 1) 2.4 1.0 1.0
assign (resid 20 and name H2' and segid 1) (resid 20 and name H2" and segid 1) 2.1 0.8 0.8
assign (resid 20 and name H2" and segid 1) (resid 20 and name H1' and segid 1) 1.8 0.7 0.7
assign (resid 20 and name H2' and segid 1) (resid 20 and name H6 and segid 1) 2.5 1.5 1.5
assign (resid 20 and name H3' and segid 1) (resid 20 and name H1' and segid 1) 2.4 2.0 2.0
assign (resid 20 and name H7* and segid 1) (resid 20 and name H1' and segid 1) 2.8 2.5 2.5
assign (resid 20 and name H7* and segid 1) (resid 20 and name H6 and segid 1) 2.3 0.9 0.9
assign (resid 21 and name H2' and segid 1) (resid 21 and name H2" and segid 1) 3.0 1.2 1.2
assign (resid 21 and name H5' and segid 1) (resid 21 and name H6 and segid 1) 3.2 2.0 2.0
assign (resid 21 and name H1' and segid 1) (resid 21 and name H6 and segid 1) 3.6 1.4 1.4
assign (resid 21 and name H2' and segid 1) (resid 21 and name H6 and segid 1) 2.2 1.5 1.5
assign (resid 21 and name H2" and segid 1) (resid 21 and name H1' and segid 1) 2.3 0.9 0.9
assign (resid 21 and name H7* and segid 1) (resid 21 and name H1' and segid 1) 3.1 2.0 2.0
assign (resid 21 and name H7* and segid 1) (resid 21 and name H6 and segid 1) 2.5 1.0 1.0
assign (resid 21 and name H2' and segid 1) (resid 21 and name H1' and segid 1) 2.1 0.8 0.8
assign (resid 21 and name H2" and segid 1) (resid 21 and name H6 and segid 1) 3.0 1.2 1.2
assign (resid 21 and name H5' and segid 1) (resid 21 and name H4' and segid 1) 1.6 1.0 1.0
assign (resid 22 and name H5 and segid 1) (resid 22 and name H6 and segid 1) 2.6 1.0 1.0
assign (resid 21 and name H7* and segid 1) (resid 22 and name H5 and segid 1) 3.0 2.5 2.5
assign (resid 22 and name H1' and segid 1) (resid 22 and name H6 and segid 1) 2.9 1.2 1.2
assign (resid 22 and name H2" and segid 1) (resid 22 and name H5 and segid 1) 3.1 2.0 2.0
assign (resid 22 and name H4' and segid 1) (resid 22 and name H6 and segid 1) 3.3 2.5 2.5
assign (resid 22 and name H2' and segid 1) (resid 22 and name H5 and segid 1) 2.7 2.0 2.0
assign (resid 22 and name H2' and segid 1) (resid 22 and name H6 and segid 1) 2.8 1.1 1.1
assign (resid 22 and name H2" and segid 1) (resid 22 and name H6 and segid 1) 2.2 1.5 1.5
assign (resid 23 and name H2' and segid 1) (resid 23 and name H1' and segid 1) 2.7 1.1 1.1
assign (resid 23 and name H2' and segid 1) (resid 23 and name H2" and segid 1) 2.3 0.9 0.9
assign (resid 23 and name H2' and segid 1) (resid 23 and name H6 and segid 1) 2.0 1.5 1.5
assign (resid 23 and name H2" and segid 1) (resid 23 and name H1' and segid 1) 2.6 1.0 1.0
assign (resid 23 and name H5 and segid 1) (resid 23 and name H6 and segid 1) 2.6 1.0 1.0
assign (resid 23 and name H1' and segid 1) (resid 23 and name H6 and segid 1) 2.7 1.1 1.1
assign (resid 23 and name H2" and segid 1) (resid 23 and name H6 and segid 1) 2.2 2.0 2.0

assign (resid 23 and name H5' and segid 1) (resid 23 and name H1' and segid 1) 4.3 1.7 1.7
assign (resid 24 and name H7* and segid 1) (resid 23 and name H1' and segid 1) 3.0 1.2 1.2
assign (resid 24 and name H7* and segid 1) (resid 23 and name H5 and segid 1) 3.0 2.0 2.0
assign (resid 24 and name H7* and segid 1) (resid 23 and name H6 and segid 1) 2.1 2.0 2.0
assign (resid 24 and name H4' and segid 1) (resid 24 and name H6 and segid 1) 4.6 1.8 1.8
assign (resid 24 and name H1' and segid 1) (resid 24 and name H6 and segid 1) 2.9 1.2 1.2
assign (resid 24 and name H4' and segid 1) (resid 24 and name H1' and segid 1) 4.9 2.0 2.0
assign (resid 24 and name H7* and segid 1) (resid 24 and name H6 and segid 1) 2.1 0.8 0.8
assign (resid 24 and name H2' and segid 1) (resid 24 and name H6 and segid 1) 1.9 1.5 1.5
assign (resid 24 and name H2" and segid 1) (resid 24 and name H6 and segid 1) 2.9 1.2 1.2
assign (resid 24 and name H3' and segid 1) (resid 24 and name H1' and segid 1) 3.5 1.4 1.4
assign (resid 24 and name H5' and segid 1) (resid 24 and name H6 and segid 1) 2.7 2.0 2.0
assign (resid 24 and name H2' and segid 1) (resid 25 and name H6 and segid 1) 3.0 1.5 1.5
assign (resid 25 and name H2" and segid 1) (resid 25 and name H6 and segid 1) 2.3 2.0 2.0
assign (resid 25 and name H4' and segid 1) (resid 25 and name H6 and segid 1) 4.1 2.1 2.1
assign (resid 25 and name H5 and segid 1) (resid 25 and name H6 and segid 1) 2.5 1.3 1.3
assign (resid 25 and name H5' and segid 1) (resid 25 and name H4' and segid 1) 1.5 1.2 1.2
assign (resid 25 and name H2' and segid 1) (resid 25 and name H1' and segid 1) 2.3 1.2 1.2
assign (resid 25 and name H3' and segid 1) (resid 25 and name H1' and segid 1) 3.5 1.8 1.8
assign (resid 25 and name H4' and segid 1) (resid 25 and name H3' and segid 1) 3.0 1.5 1.5
assign (resid 25 and name H5' and segid 1) (resid 25 and name H6 and segid 1) 4.2 2.1 2.1
assign (resid 25 and name H1' and segid 1) (resid 25 and name H6 and segid 1) 3.1 1.6 1.6
assign (resid 25 and name H2' and segid 1) (resid 25 and name H3' and segid 1) 2.6 1.3 1.3
assign (resid 25 and name H2' and segid 1) (resid 25 and name H5' and segid 1) 2.9 2.5 2.5
assign (resid 25 and name H2' and segid 1) (resid 25 and name H6 and segid 1) 2.4 1.2 1.2
assign (resid 25 and name H3' and segid 1) (resid 25 and name H6 and segid 1) 3.0 1.5 1.5
assign (resid 25 and name H5' and segid 1) (resid 25 and name H1' and segid 1) 2.9 2.0 2.0
assign (resid 25 and name H5' and segid 1) (resid 25 and name H3' and segid 1) 2.9 1.5 1.5
assign (resid 25 and name H4' and segid 1) (resid 25 and name H1' and segid 1) 3.7 1.9 1.9
assign (resid 25 and name H5" and segid 1) (resid 25 and name H4' and segid 1) 1.9 1.0 1.0
assign (resid 1 and name H1' and segid 2) (resid 1 and name H6 and segid 2) 1.8 1.2 1.2
assign (resid 1 and name H2' and segid 2) (resid 1 and name H6 and segid 2) 2.3 1.2 1.2
assign (resid 1 and name H2" and segid 2) (resid 1 and name H6 and segid 2) 2.0 1.0 1.0
assign (resid 1 and name H5 and segid 2) (resid 1 and name H6 and segid 2) 1.9 1.0 1.0
assign (resid 1 and name H5' and segid 2) (resid 1 and name H6 and segid 2) 2.6 1.3 1.3
assign (resid 2 and name H2' and segid 2) (resid 2 and name H6 and segid 2) 1.5 0.6 0.6
assign (resid 2 and name H2" and segid 2) (resid 2 and name H6 and segid 2) 2.0 0.8 0.8
assign (resid 2 and name H4' and segid 2) (resid 2 and name H6 and segid 2) 2.9 1.2 1.2
assign (resid 2 and name H5' and segid 2) (resid 2 and name H6 and segid 2) 2.5 1.0 1.0
assign (resid 2 and name H5" and segid 2) (resid 2 and name H6 and segid 2) 1.8 0.7 0.7
assign (resid 2 and name H7* and segid 2) (resid 1 and name H6 and segid 2) 2.0 0.8 0.8
assign (resid 2 and name H7* and segid 2) (resid 2 and name H6 and segid 2) 1.5 0.6 0.6
assign (resid 2 and name H7* and segid 2) (resid 3 and name H5 and segid 2) 3.1 1.1 1.1
assign (resid 3 and name H1' and segid 2) (resid 3 and name H6 and segid 2) 2.6 1.3 1.3
assign (resid 3 and name H2' and segid 2) (resid 3 and name H1' and segid 2) 1.8 1.0 1.0
assign (resid 3 and name H2' and segid 2) (resid 3 and name H6 and segid 2) 1.0 0.4 0.4
assign (resid 3 and name H2" and segid 2) (resid 3 and name H6 and segid 2) 2.5 1.3 1.3
assign (resid 3 and name H5 and segid 2) (resid 3 and name H6 and segid 2) 3.0 1.2 1.2
assign (resid 3 and name H5' and segid 2) (resid 3 and name H6 and segid 2) 2.9 1.2 1.2
assign (resid 4 and name H5 and segid 2) (resid 4 and name H6 and segid 2) 1.6 0.7 0.7
assign (resid 5 and name H2' and segid 2) (resid 5 and name H3' and segid 2) 2.1 0.8 0.8
assign (resid 5 and name H2" and segid 2) (resid 5 and name H3' and segid 2) 1.7 0.7 0.7
assign (resid 5 and name H7* and segid 2) (resid 4 and name H6 and segid 2) 2.4 1.0 1.0
assign (resid 5 and name H7* and segid 2) (resid 5 and name H6 and segid 2) 1.9 0.7 0.7
assign (resid 6 and name H1' and segid 2) (resid 6 and name H6 and segid 2) 2.7 1.1 1.1
assign (resid 6 and name H2' and segid 2) (resid 6 and name H1' and segid 2) 2.7 1.1 1.1
assign (resid 6 and name H5' and segid 2) (resid 6 and name H6 and segid 2) 2.7 1.3 1.3
assign (resid 6 and name H7* and segid 2) (resid 6 and name H6 and segid 2) 2.5 1.0 1.0
assign (resid 7 and name H2' and segid 2) (resid 7 and name H1' and segid 2) 2.4 1.0 1.0
assign (resid 7 and name H4' and segid 2) (resid 7 and name H1' and segid 2) 2.0 1.0 1.0

assign (resid 7 and name H7* and segid 2) (resid 7 and name H1' and segid 2) 3.1 1.1 1.1
assign (resid 7 and name H7* and segid 2) (resid 7 and name H6 and segid 2) 1.6 0.7 0.7
assign (resid 8 and name H2' and segid 2) (resid 8 and name H1' and segid 2) 2.0 1.0 1.0
assign (resid 8 and name H2" and segid 2) (resid 8 and name H1' and segid 2) 2.3 1.1 1.1
assign (resid 8 and name H4' and segid 2) (resid 8 and name H1' and segid 2) 2.0 0.8 0.8
assign (resid 8 and name H5 and segid 2) (resid 8 and name H6 and segid 2) 1.8 0.7 0.7
assign (resid 9 and name H2' and segid 2) (resid 9 and name H1' and segid 2) 1.6 1.0 1.0
assign (resid 9 and name H2' and segid 2) (resid 9 and name H3' and segid 2) 1.9 0.8 0.8
assign (resid 9 and name H2" and segid 2) (resid 9 and name H1' and segid 2) 1.3 1.0 1.0
assign (resid 9 and name H2" and segid 2) (resid 9 and name H3' and segid 2) 2.4 1.0 1.0
assign (resid 9 and name H4' and segid 2) (resid 9 and name H1' and segid 2) 1.5 0.6 0.6
assign (resid 9 and name H7* and segid 2) (resid 8 and name H6 and segid 2) 2.3 0.9 0.9
assign (resid 10 and name H1' and segid 2) (resid 10 and name H6 and segid 2) 2.3 0.9 0.9
assign (resid 10 and name H2' and segid 2) (resid 10 and name H1' and segid 2) 2.6 1.1 1.1
assign (resid 10 and name H7* and segid 2) (resid 10 and name H1' and segid 2) 2.8 1.2 1.2
assign (resid 10 and name H7* and segid 2) (resid 10 and name H6 and segid 2) 1.4 0.6 0.6
assign (resid 11 and name H1' and segid 2) (resid 11 and name H6 and segid 2) 2.0 1.0 1.0
assign (resid 11 and name H2' and segid 2) (resid 11 and name H1' and segid 2) 2.0 1.0 1.0
assign (resid 11 and name H2" and segid 2) (resid 11 and name H1' and segid 2) 1.7 0.7 0.7
assign (resid 11 and name H4' and segid 2) (resid 11 and name H1' and segid 2) 1.8 0.7 0.7
assign (resid 11 and name H5 and segid 2) (resid 11 and name H6 and segid 2) 2.1 0.8 0.8
assign (resid 12 and name H1' and segid 2) (resid 12 and name H6 and segid 2) 4.0 2.0 2.0
assign (resid 12 and name H2' and segid 2) (resid 12 and name H1' and segid 2) 2.2 1.1 1.1
assign (resid 12 and name H2' and segid 2) (resid 12 and name H6 and segid 2) 1.5 0.8 0.8
assign (resid 12 and name H2" and segid 2) (resid 12 and name H1' and segid 2) 1.6 0.8 0.8
assign (resid 12 and name H2" and segid 2) (resid 12 and name H3' and segid 2) 1.5 1.0 1.0
assign (resid 12 and name H5 and segid 2) (resid 12 and name H6 and segid 2) 2.5 1.3 1.3
assign (resid 12 and name H5' and segid 2) (resid 12 and name H6 and segid 2) 2.0 1.0 1.0

Exchangeable proton distance restraints file

assign (resid 3 and name H1 and segid 1) (resid 24 and name H3 and segid 1) 4.3 1.7 1.7
assign (resid 4 and name H1 and segid 1) (resid 3 and name H1 and segid 1) 4.1 1.6 1.6
assign (resid 4 and name H1 and segid 1) (resid 21 and name H3 and segid 1) 4.3 1.7 1.7
assign (resid 4 and name H1 and segid 1) (resid 5 and name H3 and segid 2) 4.3 1.7 1.7
assign (resid 5 and name H2 and segid 1) (resid 21 and name H3 and segid 1) 3.7 1.5 1.5
assign (resid 5 and name H2' and segid 1) (resid 6 and name H3 and segid 2) 4.4 1.7 1.7
assign (resid 5 and name H2" and segid 1) (resid 6 and name H3 and segid 2) 4.3 1.7 1.7
assign (resid 6 and name H2 and segid 1) (resid 20 and name H3 and segid 1) 3.5 1.4 1.4
assign (resid 6 and name H2' and segid 1) (resid 7 and name H3 and segid 2) 4.3 1.7 1.7
assign (resid 6 and name H2" and segid 1) (resid 7 and name H3 and segid 2) 4.2 1.7 1.7
assign (resid 6 and name H8 and segid 1) (resid 20 and name H3 and segid 1) 3.8 1.5 1.5
assign (resid 6 and name H61 and segid 1) (resid 6 and name H3 and segid 2) 3.8 1.5 1.5
assign (resid 7 and name H2 and segid 1) (resid 19 and name H3 and segid 1) 3.3 2.0 2.0
assign (resid 7 and name H2' and segid 1) (resid 8 and name H1 and segid 2) 5.2 2.1 2.1
assign (resid 7 and name H2" and segid 1) (resid 8 and name H1 and segid 2) 3.9 1.6 1.6
assign (resid 7 and name H61 and segid 1) (resid 7 and name H3 and segid 2) 3.9 1.6 1.6
assign (resid 8 and name H8 and segid 1) (resid 8 and name H1 and segid 2) 3.7 1.5 1.5
assign (resid 8 and name H1 and segid 1) (resid 17 and name H3 and segid 1) 4.0 1.6 1.6
assign (resid 8 and name H1 and segid 1) (resid 19 and name H3 and segid 1) 4.5 1.8 1.8
assign (resid 9 and name H2 and segid 1) (resid 17 and name H3 and segid 1) 2.9 1.2 1.2
assign (resid 9 and name H2' and segid 1) (resid 10 and name H3 and segid 2) 4.2 1.7 1.7
assign (resid 9 and name H2" and segid 1) (resid 10 and name H3 and segid 2) 4.2 1.7 1.7
assign (resid 9 and name H8 and segid 1) (resid 9 and name H3 and segid 2) 4.6 1.9 1.9
assign (resid 9 and name H61 and segid 1) (resid 9 and name H3 and segid 2) 3.7 1.5 1.5
assign (resid 10 and name H2' and segid 1) (resid 11 and name H3 and segid 2) 3.3 2.0 2.0
assign (resid 10 and name H2" and segid 1) (resid 11 and name H3 and segid 2) 3.7 1.5 1.5

assign (resid 10 and name H61 and segid 1) (resid 10 and name H3 and segid 2) 3.4 1.4 1.4
assign (resid 11 and name H8 and segid 1) (resid 11 and name H3 and segid 2) 3.4 1.4 1.4
assign (resid 11 and name H1 and segid 1) (resid 10 and name H3 and segid 2) 1.9 2.0 2.0
assign (resid 11 and name H1 and segid 1) (resid 11 and name H3 and segid 2) 4.0 1.6 1.6
assign (resid 15 and name H41 and segid 1) (resid 11 and name H41 and segid 2) 3.3 1.3 1.3
assign (resid 15 and name H41 and segid 1) (resid 11 and name H42 and segid 2) 4.4 1.8 1.8
assign (resid 15 and name H42 and segid 1) (resid 11 and name H41 and segid 2) 3.1 1.2 1.2
assign (resid 15 and name H42 and segid 1) (resid 11 and name H42 and segid 2) 3.6 1.4 1.4
assign (resid 16 and name H7* and segid 1) (resid 11 and name H1 and segid 1) 3.8 1.5 1.5
assign (resid 16 and name H7* and segid 1) (resid 16 and name H3 and segid 1) 3.7 1.5 1.5
assign (resid 17 and name H7* and segid 1) (resid 8 and name H1 and segid 1) 4.2 1.7 1.7
assign (resid 19 and name H5 and segid 1) (resid 19 and name H3 and segid 1) 4.4 1.8 1.8
assign (resid 20 and name H3 and segid 1) (resid 19 and name H3 and segid 1) 3.5 1.4 1.4
assign (resid 21 and name H7* and segid 1) (resid 21 and name H3 and segid 1) 3.9 1.6 1.6
assign (resid 22 and name H41 and segid 1) (resid 4 and name H1 and segid 1) 2.8 1.1 1.1
assign (resid 23 and name H41 and segid 1) (resid 3 and name H1 and segid 1) 2.9 1.2 1.2
assign (resid 2 and name H7* and segid 2) (resid 2 and name H3 and segid 2) 3.8 1.5 1.5
assign (resid 3 and name H3 and segid 2) (resid 3 and name H8 and segid 1) 3.9 1.6 1.6
assign (resid 4 and name H1 and segid 2) (resid 4 and name H8 and segid 1) 4.1 1.6 1.6
assign (resid 6 and name H6 and segid 2) (resid 6 and name H3 and segid 2) 4.0 1.6 1.6
assign (resid 6 and name H7* and segid 2) (resid 6 and name H3 and segid 2) 4.2 1.7 1.7
assign (resid 7 and name H3 and segid 2) (resid 6 and name H3 and segid 2) 4.6 1.9 1.9
assign (resid 7 and name H3 and segid 2) (resid 8 and name H3 and segid 2) 3.9 1.6 1.6
assign (resid 7 and name H7* and segid 2) (resid 7 and name H3 and segid 2) 3.8 1.5 1.5
assign (resid 7 and name H7* and segid 2) (resid 8 and name H3 and segid 2) 3.3 1.3 1.3
assign (resid 8 and name H41 and segid 2) (resid 8 and name H3 and segid 2) 3.9 1.6 1.6
assign (resid 8 and name H42 and segid 2) (resid 7 and name H3 and segid 2) 4.8 1.9 1.9
assign (resid 8 and name H42 and segid 2) (resid 8 and name H3 and segid 2) 4.4 1.8 1.8
assign (resid 9 and name H3 and segid 2) (resid 8 and name H3 and segid 2) 5.0 2.0 2.0
assign (resid 10 and name H6 and segid 2) (resid 10 and name H3 and segid 2) 4.1 1.6 1.6
assign (resid 10 and name H7* and segid 2) (resid 10 and name H3 and segid 2) 3.9 1.6 1.6
assign (resid 10 and name H7* and segid 2) (resid 11 and name H3 and segid 2) 3.6 1.4 1.4
assign (resid 11 and name H5 and segid 2) (resid 11 and name H3 and segid 2) 4.1 1.7 1.7
assign (resid 11 and name H41 and segid 2) (resid 10 and name H3 and segid 2) 4.8 1.9 1.9
assign (resid 11 and name H41 and segid 2) (resid 11 and name H3 and segid 2) 4.4 1.8 1.8
assign (resid 11 and name H42 and segid 2) (resid 10 and name H3 and segid 2) 4.0 1.6 1.6
assign (resid 11 and name H42 and segid 2) (resid 11 and name H3 and segid 2) 4.1 1.7 1.7

A3. NOe distance restraints files for non-modified triplex

Non-exchangeable proton distance restraints file

assign (residue 1 and name H2' and segid 1) (residue 1 and name H1' and segid 1) 2.4 1.7 1.7
assign (residue 1 and name H2' and segid 1) (residue 1 and name H8 and segid 1) 2.9 1.8 1.8
assign (residue 1 and name H2" and segid 1) (residue 1 and name H4' and segid 1) 2.5 1.5 1.5
assign (residue 1 and name H2" and segid 1) (residue 1 and name H1' and segid 1) 3.3 1.7 1.7
assign (residue 1 and name H1' and segid 1) (residue 1 and name H8 and segid 1) 2.1 1.5 1.5
assign (residue 1 and name H2' and segid 1) (residue 1 and name H2" and segid 1) 2.0 1.0 1.0
assign (residue 1 and name H2" and segid 1) (residue 1 and name H8 and segid 1) 3.2 1.6 1.6
assign (residue 2 and name H1' and segid 1) (residue 2 and name H8 and segid 1) 3.9 1.5 1.5
assign (residue 2 and name H8 and segid 1) (residue 1 and name H8 and segid 1) 4.4 2.2 2.2
assign (residue 2 and name H2' and segid 1) (residue 1 and name H8 and segid 1) 3.0 1.5 1.5

assign (residue 2 and name H1' and segid 1) (residue 1 and name H8 and segid 1) 3.1 2.0 2.0
assign (residue 2 and name H2' and segid 1) (residue 2 and name H1' and segid 1) 2.4 1.5 1.5
assign (residue 2 and name H2' and segid 1) (residue 2 and name H2" and segid 1) 2.2 0.9 0.9
assign (residue 2 and name H2' and segid 1) (residue 2 and name H8 and segid 1) 2.4 2.0 2.0
assign (residue 2 and name H2" and segid 1) (residue 2 and name H1' and segid 1) 1.6 0.6 0.6
assign (residue 2 and name H3' and segid 1) (residue 2 and name H8 and segid 1) 3.0 2.0 2.0
assign (residue 2 and name H5' and segid 1) (residue 2 and name H8 and segid 1) 3.8 1.5 1.5
assign (residue 2 and name H5" and segid 1) (residue 2 and name H8 and segid 1) 3.9 2.5 2.5
assign (residue 2 and name H2' and segid 1) (residue 2 and name H4' and segid 1) 4.6 1.8 1.8
assign (residue 2 and name H2" and segid 1) (residue 2 and name H8 and segid 1) 2.4 1.0 1.0
assign (residue 2 and name H4' and segid 1) (residue 2 and name H8 and segid 1) 4.8 1.9 1.9
assign (residue 3 and name H1' and segid 1) (residue 2 and name H8 and segid 1) 3.8 2.5 2.5
assign (residue 3 and name H2' and segid 1) (residue 2 and name H8 and segid 1) 2.9 2.0 2.0
assign (residue 2 and name H8 and segid 1) (residue 3 and name H8 and segid 1) 3.9 2.5 2.5
assign (residue 3 and name H2' and segid 1) (residue 3 and name H1' and segid 1) 2.1 1.0 1.0
assign (residue 3 and name H1' and segid 1) (residue 3 and name H8 and segid 1) 3.2 1.3 1.3
assign (residue 3 and name H2' and segid 1) (residue 3 and name H2" and segid 1) 2.5 1.0 1.0
assign (residue 3 and name H2' and segid 1) (residue 3 and name H8 and segid 1) 2.4 1.5 1.5
assign (residue 3 and name H2" and segid 1) (residue 3 and name H1' and segid 1) 2.9 1.2 1.2
assign (residue 3 and name H2' and segid 1) (residue 3 and name H4' and segid 1) 3.4 1.4 1.4
assign (residue 3 and name H2" and segid 1) (residue 3 and name H8 and segid 1) 2.8 1.1 1.1
assign (residue 3 and name H3' and segid 1) (residue 3 and name H1' and segid 1) 2.9 1.2 1.2
assign (residue 3 and name H4' and segid 1) (residue 3 and name H8 and segid 1) 3.9 1.6 1.6
assign (residue 3 and name H3' and segid 1) (residue 3 and name H8 and segid 1) 3.1 1.8 1.8
assign (residue 3 and name H8 and segid 1) (residue 4 and name H8 and segid 1) 4.0 3.0 3.0
assign (residue 4 and name H2' and segid 1) (residue 3 and name H8 and segid 1) 3.7 1.5 1.5
assign (residue 4 and name H2' and segid 1) (residue 4 and name H2" and segid 1) 2.2 0.9 0.9
assign (residue 4 and name H2' and segid 1) (residue 4 and name H3' and segid 1) 2.5 1.0 1.0
assign (residue 4 and name H2' and segid 1) (residue 4 and name H4' and segid 1) 3.3 1.3 1.3
assign (residue 4 and name H2' and segid 1) (residue 4 and name H8 and segid 1) 4.1 1.6 1.6
assign (residue 4 and name H2" and segid 1) (residue 4 and name H1' and segid 1) 2.3 0.9 0.9
assign (residue 4 and name H2" and segid 1) (residue 4 and name H3' and segid 1) 2.2 0.9 0.9
assign (residue 4 and name H2" and segid 1) (residue 4 and name H4' and segid 1) 2.7 1.5 1.5
assign (residue 4 and name H5' and segid 1) (residue 4 and name H4' and segid 1) 2.0 0.8 0.8
assign (residue 4 and name H1' and segid 1) (residue 4 and name H8 and segid 1) 3.1 1.3 1.3
assign (residue 4 and name H3' and segid 1) (residue 4 and name H1' and segid 1) 2.3 1.5 1.5
assign (residue 4 and name H4' and segid 1) (residue 4 and name H8 and segid 1) 3.7 2.0 2.0
assign (residue 4 and name H2' and segid 1) (residue 4 and name H1' and segid 1) 2.5 1.0 1.0
assign (residue 4 and name H2" and segid 1) (residue 4 and name H8 and segid 1) 2.3 0.9 0.9
assign (residue 4 and name H4' and segid 1) (residue 4 and name H1' and segid 1) 3.1 1.2 1.2
assign (residue 4 and name H3' and segid 1) (residue 4 and name H8 and segid 1) 3.5 2.0 2.0
assign (residue 5 and name H8 and segid 1) (residue 4 and name H8 and segid 1) 4.5 1.8 1.8
assign (residue 5 and name H1' and segid 1) (residue 4 and name H8 and segid 1) 3.4 2.3 2.3
assign (residue 5 and name H2' and segid 1) (residue 4 and name H8 and segid 1) 2.5 1.0 1.0
assign (residue 5 and name H2" and segid 1) (residue 4 and name H8 and segid 1) 3.2 2.0 2.0
assign (residue 5 and name H2' and segid 1) (residue 5 and name H1' and segid 1) 2.9 1.2 1.2
assign (residue 5 and name H2" and segid 1) (residue 5 and name H1' and segid 1) 2.6 1.0 1.0
assign (residue 5 and name H1' and segid 1) (residue 5 and name H8 and segid 1) 1.9 1.2 1.2
assign (residue 5 and name H2' and segid 1) (residue 5 and name H2" and segid 1) 2.4 1.0 1.0
assign (residue 5 and name H2' and segid 1) (residue 5 and name H8 and segid 1) 3.2 1.3 1.3
assign (residue 5 and name H8 and segid 1) (residue 6 and name H8 and segid 1) 3.8 1.5 1.5
assign (residue 6 and name H2" and segid 1) (residue 5 and name H8 and segid 1) 3.5 2.0 2.0
assign (residue 6 and name H2" and segid 1) (residue 6 and name H8 and segid 1) 2.2 0.9 0.9
assign (residue 6 and name H1' and segid 1) (residue 5 and name H8 and segid 1) 3.7 2.0 2.0
assign (residue 6 and name H2' and segid 1) (residue 5 and name H8 and segid 1) 2.6 1.8 1.8
assign (residue 6 and name H2' and segid 1) (residue 6 and name H2" and segid 1) 2.7 1.1 1.1
assign (residue 6 and name H2' and segid 1) (residue 6 and name H3' and segid 1) 3.3 1.3 1.3
assign (residue 6 and name H2' and segid 1) (residue 6 and name H8 and segid 1) 3.0 1.2 1.2
assign (residue 6 and name H2" and segid 1) (residue 6 and name H1' and segid 1) 3.3 1.3 1.3
assign (residue 6 and name H2' and segid 1) (residue 6 and name H1' and segid 1) 3.1 1.2 1.2

assign (residue 6 and name H4' and segid 1) (residue 6 and name H3' and segid 1) 2.7 1.1 1.1
assign (residue 6 and name H2' and segid 1) (residue 6 and name H4' and segid 1) 3.5 1.4 1.4
assign (residue 6 and name H3' and segid 1) (residue 6 and name H8 and segid 1) 3.1 2.0 2.0
assign (residue 6 and name H4' and segid 1) (residue 6 and name H1' and segid 1) 2.4 1.0 1.0
assign (residue 6 and name H1' and segid 1) (residue 6 and name H8 and segid 1) 2.6 1.0 1.0
assign (residue 6 and name H3' and segid 1) (residue 6 and name H1' and segid 1) 3.1 1.2 1.2
assign (residue 6 and name H4' and segid 1) (residue 6 and name H8 and segid 1) 3.7 2.0 2.0
assign (residue 6 and name H8 and segid 1) (residue 7 and name H8 and segid 1) 4.0 2.0 2.0
assign (residue 7 and name H1' and segid 1) (residue 6 and name H8 and segid 1) 3.7 3.2 3.2
assign (residue 7 and name H1' and segid 1) (residue 7 and name H8 and segid 1) 2.3 1.5 1.5
assign (residue 7 and name H2' and segid 1) (residue 7 and name H2" and segid 1) 2.1 0.8 0.8
assign (residue 7 and name H2" and segid 1) (residue 7 and name H1' and segid 1) 2.2 0.9 0.9
assign (residue 7 and name H2" and segid 1) (residue 7 and name H8 and segid 1) 4.7 1.9 1.9
assign (residue 7 and name H4' and segid 1) (residue 7 and name H1' and segid 1) 4 1.6 1.6
assign (residue 7 and name H4' and segid 1) (residue 7 and name H3' and segid 1) 2.5 1.0 1.0
assign (residue 7 and name H3' and segid 1) (residue 7 and name H8 and segid 1) 3.7 2.0 2.0
assign (residue 7 and name H5" and segid 1) (residue 7 and name H3' and segid 1) 2.9 1.2 1.2
assign (residue 7 and name H3' and segid 1) (residue 7 and name H1' and segid 1) 3.1 1.2 1.2
assign (residue 7 and name H2' and segid 1) (residue 7 and name H1' and segid 1) 2.2 1.3 1.3
assign (residue 7 and name H2' and segid 1) (residue 7 and name H8 and segid 1) 2.3 1.0 1.0
assign (residue 7 and name H3' and segid 1) (residue 7 and name H1' and segid 1) 4.0 1.6 1.6
assign (residue 8 and name H2' and segid 1) (residue 8 and name H1' and segid 1) 2.1 0.9 0.9
assign (residue 8 and name H2' and segid 1) (residue 8 and name H2" and segid 1) 2.9 1.2 1.2
assign (residue 8 and name H2' and segid 1) (residue 8 and name H8 and segid 1) 3.0 1.0 1.0
assign (residue 8 and name H2" and segid 1) (residue 8 and name H8 and segid 1) 3.1 1.2 1.2
assign (residue 8 and name H1' and segid 1) (residue 8 and name H8 and segid 1) 3.3 1.3 1.3
assign (residue 8 and name H2" and segid 1) (residue 8 and name H1' and segid 1) 2.3 1.0 1.0
assign (residue 8 and name H5' and segid 1) (residue 8 and name H8 and segid 1) 3.1 2.0 2.0
assign (residue 9 and name H8 and segid 1) (residue 8 and name H8 and segid 1) 3.3 1.3 1.3
assign (residue 9 and name H2' and segid 1) (residue 8 and name H8 and segid 1) 3.2 2.5 2.5
assign (residue 9 and name H2' and segid 1) (residue 9 and name H1' and segid 1) 3.3 1.3 1.3
assign (residue 9 and name H2' and segid 1) (residue 9 and name H2" and segid 1) 2.0 0.8 0.8
assign (residue 9 and name H2" and segid 1) (residue 9 and name H1' and segid 1) 2.7 1.1 1.1
assign (residue 9 and name H2" and segid 1) (residue 9 and name H8 and segid 1) 2.7 1.5 1.5
assign (residue 9 and name H1' and segid 1) (residue 9 and name H8 and segid 1) 3.1 1.2 1.2
assign (residue 10 and name H2" and segid 1) (residue 10 and name H8 and segid 1) 2.9 1.2 1.2
assign (residue 10 and name H1' and segid 1) (residue 10 and name H8 and segid 1) 4.6 1.8 1.8
assign (residue 10 and name H2' and segid 1) (residue 10 and name H3' and segid 1) 2.2 0.9 0.9
assign (residue 10 and name H2" and segid 1) (residue 10 and name H1' and segid 1) 2.4 1.0 1.0
assign (residue 10 and name H2" and segid 1) (residue 10 and name H3' and segid 1) 2.1 0.8 0.8
assign (residue 10 and name H2' and segid 1) (residue 10 and name H1' and segid 1) 4.2 1.7 1.7
assign (residue 10 and name H3' and segid 1) (residue 10 and name H8 and segid 1) 2.8 2.0 2.0
assign (residue 10 and name H3' and segid 1) (residue 10 and name H1' and segid 1) 4.3 1.7 1.7
assign (residue 11 and name H8 and segid 1) (residue 10 and name H8 and segid 1) 4.0 1.6 1.6
assign (residue 11 and name H1' and segid 1) (residue 10 and name H8 and segid 1) 3.5 2.5 2.5
assign (residue 11 and name H2' and segid 1) (residue 10 and name H8 and segid 1) 3.2 2.0 2.0
assign (residue 11 and name H2' and segid 1) (residue 11 and name H2" and segid 1) 2.3 0.9 0.9
assign (residue 11 and name H2' and segid 1) (residue 11 and name H8 and segid 1) 2.7 1.1 1.1
assign (residue 11 and name H2" and segid 1) (residue 11 and name H1' and segid 1) 1.8 0.7 0.7
assign (residue 11 and name H2" and segid 1) (residue 11 and name H8 and segid 1) 3.0 1.2 1.2
assign (residue 11 and name H1' and segid 1) (residue 11 and name H8 and segid 1) 3.3 1.3 1.3
assign (residue 11 and name H2' and segid 1) (residue 11 and name H1' and segid 1) 4.2 1.7 1.7
assign (residue 12 and name H2' and segid 1) (residue 11 and name H8 and segid 1) 3.4 1.7 1.7
assign (residue 12 and name H2" and segid 1) (residue 11 and name H8 and segid 1) 3.2 1.6 1.6
assign (residue 12 and name H2' and segid 1) (residue 12 and name H5' and segid 1) 2.1 1.2 1.2
assign (residue 12 and name H2" and segid 1) (residue 12 and name H1' and segid 1) 2.3 1.2 1.2
assign (residue 12 and name H5' and segid 1) (residue 12 and name H1' and segid 1) 3.2 2.0 2.0
assign (residue 12 and name H2' and segid 1) (residue 12 and name H8 and segid 1) 4.9 2.0 2.0
assign (residue 12 and name H2" and segid 1) (residue 12 and name H8 and segid 1) 4.0 1.7 1.7
assign (residue 12 and name H2' and segid 1) (residue 12 and name H4' and segid 1) 4.1 2.1 2.1

assign (residue 12 and name H1' and segid 1) (residue 12 and name H8 and segid 1) 4.4 1.8 1.8
assign (residue 12 and name H2' and segid 1) (residue 12 and name H2" and segid 1) 2.7 1.4 1.4
assign (residue 12 and name H4' and segid 1) (residue 12 and name H1' and segid 1) 3.0 1.2 1.2
assign (residue 14 and name H1' and segid 1) (residue 14 and name H6 and segid 1) 3.3 1.3 1.3
assign (residue 14 and name H2' and segid 1) (residue 14 and name H5" and segid 1) 3.1 1.6 1.6
assign (residue 14 and name H2" and segid 1) (residue 14 and name H1' and segid 1) 3.7 1.7 1.7
assign (residue 14 and name H2" and segid 1) (residue 14 and name H4' and segid 1) 2.6 1.3 1.3
assign (residue 14 and name H5 and segid 1) (residue 14 and name H6 and segid 1) 2.9 1.1 1.1
assign (residue 14 and name H2' and segid 1) (residue 14 and name H2" and segid 1) 1.6 0.8 0.8
assign (residue 14 and name H2' and segid 1) (residue 14 and name H6 and segid 1) 1.8 0.9 0.9
assign (residue 14 and name H4' and segid 1) (residue 14 and name H1' and segid 1) 3.7 1.5 1.5
assign (residue 14 and name H4' and segid 1) (residue 14 and name H6 and segid 1) 3.3 2.0 2.0
assign (residue 14 and name H2' and segid 1) (residue 14 and name H5' and segid 1) 3.3 1.7 1.7
assign (residue 14 and name H5' and segid 1) (residue 14 and name H1' and segid 1) 2.9 1.5 1.5
assign (residue 14 and name H2' and segid 1) (residue 14 and name H1' and segid 1) 4.0 2.0 2.0
assign (residue 15 and name H1' and segid 1) (residue 15 and name H6 and segid 1) 3.1 1.2 1.2
assign (residue 15 and name H2' and segid 1) (residue 15 and name H1' and segid 1) 2.6 1.0 1.0
assign (residue 15 and name H2' and segid 1) (residue 15 and name H2" and segid 1) 2.1 0.8 0.8
assign (residue 15 and name H2" and segid 1) (residue 15 and name H1' and segid 1) 1.4 1.0 1.0
assign (residue 15 and name H4' and segid 1) (residue 15 and name H6 and segid 1) 2.9 1.2 1.2
assign (residue 15 and name H5 and segid 1) (residue 15 and name H6 and segid 1) 2.3 1.0 1.0
assign (residue 15 and name H5" and segid 1) (residue 15 and name H6 and segid 1) 2.4 1.8 1.8
assign (residue 15 and name H5' and segid 1) (residue 15 and name H6 and segid 1) 2.2 1.1 1.1
assign (residue 15 and name H2' and segid 1) (residue 15 and name H6 and segid 1) 2.6 1.5 1.5
assign (residue 15 and name H2" and segid 1) (residue 15 and name H6 and segid 1) 2.4 1.2 1.2
assign (residue 15 and name H2" and segid 1) (residue 16 and name H6 and segid 1) 2.4 1.0 1.0
assign (residue 16 and name H7* and segid 1) (residue 15 and name H1' and segid 1) 2.8 1.1 1.1
assign (residue 16 and name H7* and segid 1) (residue 15 and name H6 and segid 1) 2.6 1.0 1.0
assign (residue 16 and name H7* and segid 1) (residue 15 and name H5 and segid 1) 3.0 2.0 2.0
assign (residue 16 and name H1' and segid 1) (residue 16 and name H6 and segid 1) 4.6 1.8 1.8
assign (residue 16 and name H2' and segid 1) (residue 16 and name H2" and segid 1) 3.0 1.2 1.2
assign (residue 16 and name H2' and segid 1) (residue 16 and name H6 and segid 1) 2.4 1.8 1.8
assign (residue 16 and name H2" and segid 1) (residue 16 and name H6 and segid 1) 2.9 1.5 1.5
assign (residue 16 and name H7* and segid 1) (residue 16 and name H1' and segid 1) 3.1 2.0 2.0
assign (residue 16 and name H7* and segid 1) (residue 16 and name H6 and segid 1) 2.1 0.8 0.8
assign (residue 16 and name H2' and segid 1) (residue 16 and name H1' and segid 1) 1.9 1.0 1.0
assign (residue 16 and name H2" and segid 1) (residue 16 and name H1' and segid 1) 3.1 1.2 1.2
assign (residue 16 and name H2" and segid 1) (residue 17 and name H6 and segid 1) 3.8 1.5 1.5
assign (residue 17 and name H7* and segid 1) (residue 17 and name H6 and segid 1) 2.9 1.1 1.1
assign (residue 17 and name H2' and segid 1) (residue 17 and name H1' and segid 1) 2.4 1.0 1.0
assign (residue 17 and name H2' and segid 1) (residue 17 and name H2" and segid 1) 1.7 0.7 0.7
assign (residue 17 and name H1' and segid 1) (residue 17 and name H6 and segid 1) 3.3 1.3 1.3
assign (residue 17 and name H2" and segid 1) (residue 17 and name H1' and segid 1) 2.1 0.9 0.9
assign (residue 17 and name H7* and segid 1) (residue 18 and name H5 and segid 1) 3.4 1.4 1.4
assign (residue 17 and name H4' and segid 1) (residue 17 and name H1' and segid 1) 2.4 1.0 1.0
assign (residue 18 and name H2' and segid 1) (residue 18 and name H2" and segid 1) 2.8 1.1 1.1
assign (residue 18 and name H1' and segid 1) (residue 18 and name H6 and segid 1) 3.9 1.6 1.6
assign (residue 18 and name H2' and segid 1) (residue 18 and name H6 and segid 1) 3.9 2.0 2.0
assign (residue 18 and name H2" and segid 1) (residue 18 and name H6 and segid 1) 3.1 1.2 1.2
assign (residue 18 and name H5 and segid 1) (residue 18 and name H6 and segid 1) 2.0 0.8 0.8
assign (residue 18 and name H2' and segid 1) (residue 18 and name H1' and segid 1) 2.1 1.4 1.4
assign (residue 18 and name H2" and segid 1) (residue 18 and name H1' and segid 1) 2.0 1.0 1.0
assign (residue 18 and name H3' and segid 1) (residue 18 and name H6 and segid 1) 3.3 1.3 1.3
assign (residue 18 and name H5' and segid 1) (residue 18 and name H6 and segid 1) 2.8 1.8 1.8
assign (residue 19 and name H1' and segid 1) (residue 19 and name H6 and segid 1) 2.7 1.1 1.1
assign (residue 19 and name H2" and segid 1) (residue 19 and name H6 and segid 1) 2.4 1.5 1.5
assign (residue 19 and name H2" and segid 1) (residue 19 and name H1' and segid 1) 2.7 1.1 1.1
assign (residue 19 and name H5 and segid 1) (residue 19 and name H6 and segid 1) 2.3 0.9 0.9
assign (residue 19 and name H2' and segid 1) (residue 19 and name H1' and segid 1) 2.7 1.1 1.1
assign (residue 19 and name H2' and segid 1) (residue 19 and name H6 and segid 1) 2.7 1.1 1.1

assign (residue 19 and name H5' and segid 1) (residue 19 and name H6 and segid 1) 4.1 1.6 1.6
assign (residue 20 and name H7* and segid 1) (residue 19 and name H6 and segid 1) 5.0 2.0 2.0
assign (residue 20 and name H7* and segid 1) (residue 19 and name H5 and segid 1) 3.6 2.0 2.0
assign (residue 20 and name H1' and segid 1) (residue 20 and name H6 and segid 1) 3.2 1.3 1.3
assign (residue 20 and name H2' and segid 1) (residue 20 and name H1' and segid 1) 2.6 1.1 1.1
assign (residue 20 and name H2' and segid 1) (residue 20 and name H2" and segid 1) 2.1 0.8 0.8
assign (residue 20 and name H2" and segid 1) (residue 20 and name H1' and segid 1) 2.3 1.0 1.0
assign (residue 20 and name H2" and segid 1) (residue 20 and name H6 and segid 1) 2.5 1.0 1.0
assign (residue 20 and name H2' and segid 1) (residue 20 and name H6 and segid 1) 2.8 1.1 1.1
assign (residue 20 and name H3' and segid 1) (residue 20 and name H1' and segid 1) 2.4 2.0 2.0
assign (residue 20 and name H7* and segid 1) (residue 20 and name H1' and segid 1) 2.8 2.5 2.5
assign (residue 20 and name H7* and segid 1) (residue 20 and name H6 and segid 1) 2.8 1.1 1.1
assign (residue 21 and name H2' and segid 1) (residue 21 and name H2" and segid 1) 3.0 1.2 1.2
assign (residue 21 and name H5' and segid 1) (residue 21 and name H1' and segid 1) 3.3 1.3 1.3
assign (residue 21 and name H1' and segid 1) (residue 21 and name H6 and segid 1) 2.6 1.0 1.0
assign (residue 21 and name H2' and segid 1) (residue 21 and name H6 and segid 1) 2.2 1.5 1.5
assign (residue 21 and name H2" and segid 1) (residue 21 and name H1' and segid 1) 2.6 1.4 1.4
assign (residue 21 and name H7* and segid 1) (residue 21 and name H1' and segid 1) 3.1 2.0 2.0
assign (residue 21 and name H7* and segid 1) (residue 21 and name H6 and segid 1) 2.0 0.8 0.8
assign (residue 21 and name H2' and segid 1) (residue 21 and name H1' and segid 1) 2.4 1.0 1.0
assign (residue 21 and name H2" and segid 1) (residue 21 and name H6 and segid 1) 3.0 1.2 1.2
assign (residue 21 and name H3' and segid 1) (residue 21 and name H1' and segid 1) 2.6 1.5 1.5
assign (residue 22 and name H5 and segid 1) (residue 22 and name H6 and segid 1) 2.3 1.0 1.0
assign (residue 21 and name H7* and segid 1) (residue 22 and name H1' and segid 1) 3.1 2.0 2.0
assign (residue 22 and name H5 and segid 1) (residue 21 and name H6 and segid 1) 3.0 1.5 1.5
assign (residue 22 and name H1' and segid 1) (residue 22 and name H6 and segid 1) 2.9 1.2 1.2
assign (residue 22 and name H2" and segid 1) (residue 22 and name H1' and segid 1) 1.8 1.0 1.0
assign (residue 22 and name H4' and segid 1) (residue 22 and name H6 and segid 1) 3.3 2.5 2.5
assign (residue 22 and name H2' and segid 1) (residue 22 and name H5 and segid 1) 2.7 2.0 2.0
assign (residue 22 and name H2' and segid 1) (residue 22 and name H6 and segid 1) 2.8 1.1 1.1
assign (residue 22 and name H2" and segid 1) (residue 22 and name H6 and segid 1) 2.5 1.0 1.0
assign (residue 23 and name H2' and segid 1) (residue 23 and name H1' and segid 1) 2.7 1.1 1.1
assign (residue 23 and name H2' and segid 1) (residue 23 and name H2" and segid 1) 2.3 0.9 0.9
assign (residue 23 and name H2' and segid 1) (residue 23 and name H6 and segid 1) 2.0 1.5 1.5
assign (residue 23 and name H2" and segid 1) (residue 23 and name H1' and segid 1) 2.6 1.0 1.0
assign (residue 23 and name H5 and segid 1) (residue 23 and name H6 and segid 1) 2.4 0.9 0.9
assign (residue 23 and name H1' and segid 1) (residue 23 and name H6 and segid 1) 2.2 0.9 0.9
assign (residue 23 and name H2" and segid 1) (residue 23 and name H6 and segid 1) 2.2 0.9 0.9
assign (residue 23 and name H3' and segid 1) (residue 23 and name H6 and segid 1) 2.7 1.8 1.8
assign (residue 24 and name H7* and segid 1) (residue 23 and name H6 and segid 1) 2.7 1.1 1.1
assign (residue 24 and name H7* and segid 1) (residue 23 and name H1' and segid 1) 3.0 2.0 2.0
assign (residue 24 and name H7* and segid 1) (residue 23 and name H5 and segid 1) 3.0 2.0 2.0
assign (residue 24 and name H2" and segid 1) (residue 23 and name H6 and segid 1) 3.6 1.4 1.4
assign (residue 24 and name H1' and segid 1) (residue 24 and name H6 and segid 1) 2.5 1.3 1.3
assign (residue 24 and name H7* and segid 1) (residue 24 and name H6 and segid 1) 2.7 1.2 1.2
assign (residue 24 and name H2' and segid 1) (residue 24 and name H6 and segid 1) 1.9 1.5 1.5
assign (residue 24 and name H2' and segid 1) (residue 24 and name H1' and segid 1) 2.5 1.0 1.0
assign (residue 24 and name H2" and segid 1) (residue 24 and name H6 and segid 1) 2.7 1.1 1.1
assign (residue 24 and name H2' and segid 1) (residue 25 and name H6 and segid 1) 3.0 1.5 1.5
assign (residue 24 and name H2" and segid 1) (residue 24 and name H1' and segid 1) 2.2 0.9 0.9
assign (residue 25 and name H5 and segid 1) (residue 24 and name H6 and segid 1) 3.3 1.3 1.3
assign (residue 25 and name H2" and segid 1) (residue 25 and name H6 and segid 1) 3.3 1.3 1.3
assign (residue 25 and name H4' and segid 1) (residue 25 and name H6 and segid 1) 3.7 1.5 1.5
assign (residue 25 and name H5 and segid 1) (residue 25 and name H6 and segid 1) 2.6 1.0 1.0
assign (residue 25 and name H5" and segid 1) (residue 25 and name H6 and segid 1) 2.9 1.2 1.2
assign (residue 25 and name H2' and segid 1) (residue 25 and name H1' and segid 1) 2.2 1.1 1.1
assign (residue 25 and name H3' and segid 1) (residue 25 and name H1' and segid 1) 2.6 1.3 1.3
assign (residue 25 and name H5' and segid 1) (residue 25 and name H6 and segid 1) 3.3 1.3 1.3
assign (residue 25 and name H1' and segid 1) (residue 25 and name H6 and segid 1) 3.1 1.6 1.6
assign (residue 25 and name H2' and segid 1) (residue 25 and name H6 and segid 1) 2.1 0.9 0.9

assign (residue 25 and name H2" and segid 1) (residue 25 and name H1' and segid 1) 2.7 1.1 1.1
assign (residue 25 and name H2' and segid 1) (residue 25 and name H6 and segid 1) 2.4 1.2 1.2
assign (residue 25 and name H3' and segid 1) (residue 25 and name H6 and segid 1) 3.0 1.5 1.5
assign (residue 25 and name H5' and segid 1) (residue 25 and name H1' and segid 1) 2.9 2.0 2.0
assign (residue 25 and name H5' and segid 1) (residue 25 and name H3' and segid 1) 2.9 1.5 1.5
assign (residue 25 and name H4' and segid 1) (residue 25 and name H1' and segid 1) 2.3 0.9 0.9
assign (residue 1 and name H1' and segid 2) (residue 1 and name H6 and segid 2) 3.2 1.3 1.3
assign (residue 1 and name H2' and segid 2) (residue 1 and name H1' and segid 2) 2.3 1.2 1.2
assign (residue 1 and name H2' and segid 2) (residue 1 and name H6 and segid 2) 2.2 1.1 1.1
assign (residue 1 and name H2" and segid 2) (residue 1 and name H1' and segid 2) 2.8 1.1 1.1
assign (residue 1 and name H2" and segid 2) (residue 1 and name H6 and segid 2) 3.0 1.5 1.5
assign (residue 1 and name H3' and segid 2) (residue 1 and name H1' and segid 2) 3.0 1.5 1.5
assign (residue 1 and name H4' and segid 2) (residue 1 and name H1' and segid 2) 2.6 1.0 1.0
assign (residue 1 and name H5 and segid 2) (residue 1 and name H6 and segid 2) 2.5 1.0 1.0
assign (residue 2 and name H7* and segid 2) (residue 1 and name H6 and segid 2) 2.8 1.1 1.1
assign (residue 2 and name H7* and segid 2) (residue 1 and name H5 and segid 2) 2.9 1.2 1.2
assign (residue 2 and name H1' and segid 2) (residue 2 and name H6 and segid 2) 3.0 1.2 1.2
assign (residue 2 and name H2' and segid 2) (residue 2 and name H6 and segid 2) 2.0 1.0 1.0
assign (residue 2 and name H2" and segid 2) (residue 2 and name H6 and segid 2) 3.0 1.2 1.2
assign (residue 2 and name H4' and segid 2) (residue 2 and name H6 and segid 2) 2.9 1.2 1.2
assign (residue 2 and name H5' and segid 2) (residue 2 and name H6 and segid 2) 2.5 1.0 1.0
assign (residue 2 and name H5" and segid 2) (residue 2 and name H6 and segid 2) 1.8 1.0 1.0
assign (residue 2 and name H7* and segid 2) (residue 2 and name H6 and segid 2) 2.5 1.2 1.2
assign (residue 3 and name H5 and segid 2) (residue 2 and name H6 and segid 2) 3.5 2.0 2.0
assign (residue 3 and name H1' and segid 2) (residue 3 and name H6 and segid 2) 2.6 1.1 1.1
assign (residue 3 and name H2" and segid 2) (residue 3 and name H1' and segid 2) 2.7 1.1 1.1
assign (residue 3 and name H2' and segid 2) (residue 3 and name H6 and segid 2) 1.8 1.0 1.0
assign (residue 3 and name H2" and segid 2) (residue 3 and name H6 and segid 2) 2.5 1.3 1.3
assign (residue 3 and name H5 and segid 2) (residue 3 and name H6 and segid 2) 2.8 1.1 1.1
assign (residue 3 and name H5" and segid 2) (residue 3 and name H1' and segid 2) 3.1 2.0 2.0
assign (residue 3 and name H3' and segid 2) (residue 3 and name H1' and segid 2) 2.6 1.5 1.5
assign (residue 3 and name H4' and segid 2) (residue 3 and name H1' and segid 2) 2.8 1.1 1.1
assign (residue 4 and name H5 and segid 2) (residue 4 and name H6 and segid 2) 3.6 1.4 1.4
assign (residue 4 and name H1' and segid 2) (residue 4 and name H6 and segid 2) 3.6 1.4 1.4
assign (residue 4 and name H2' and segid 2) (residue 4 and name H1' and segid 2) 3.4 1.3 1.3
assign (residue 4 and name H2" and segid 2) (residue 4 and name H1' and segid 2) 3.0 1.2 1.2
assign (residue 5 and name H1' and segid 2) (residue 5 and name H6 and segid 2) 3.4 1.4 1.4
assign (residue 5 and name H2' and segid 2) (residue 5 and name H1' and segid 2) 2.2 0.9 0.9
assign (residue 5 and name H2" and segid 2) (residue 5 and name H1' and segid 2) 2.1 0.9 0.9
assign (residue 5 and name H2" and segid 2) (residue 5 and name H6 and segid 2) 2.2 1.5 1.5
assign (residue 5 and name H3' and segid 2) (residue 5 and name H6 and segid 2) 3.6 1.4 1.4
assign (residue 5 and name H5' and segid 2) (residue 5 and name H6 and segid 2) 4.0 1.6 1.6
assign (residue 5 and name H7* and segid 2) (residue 5 and name H6 and segid 2) 3.4 2.0 2.0
assign (residue 6 and name H1' and segid 2) (residue 6 and name H6 and segid 2) 3.0 1.2 1.2
assign (residue 6 and name H2' and segid 2) (residue 6 and name H1' and segid 2) 3.5 1.4 1.4
assign (residue 6 and name H2' and segid 2) (residue 6 and name H6 and segid 2) 4.3 2.0 2.0
assign (residue 6 and name H2" and segid 2) (residue 6 and name H1' and segid 2) 2.4 1.5 1.5
assign (residue 6 and name H2" and segid 2) (residue 6 and name H6 and segid 2) 1.6 0.8 0.8
assign (residue 6 and name H3' and segid 2) (residue 7 and name H6 and segid 2) 3.9 2.0 2.0
assign (residue 6 and name H4' and segid 2) (residue 6 and name H6 and segid 2) 3.4 1.8 1.8
assign (residue 6 and name H7* and segid 2) (residue 5 and name H6 and segid 2) 3.2 2.0 2.0
assign (residue 6 and name H7* and segid 2) (residue 6 and name H6 and segid 2) 2.4 1.0 1.0
assign (residue 7 and name H1' and segid 2) (residue 7 and name H6 and segid 2) 2.7 1.1 1.1
assign (residue 7 and name H2' and segid 2) (residue 6 and name H6 and segid 2) 2.6 1.5 1.5
assign (residue 7 and name H2' and segid 2) (residue 7 and name H1' and segid 2) 2.2 1.2 1.2
assign (residue 7 and name H2" and segid 2) (residue 7 and name H1' and segid 2) 1.5 0.6 0.6
assign (residue 7 and name H2" and segid 2) (residue 7 and name H6 and segid 2) 2.5 1.8 1.8
assign (residue 7 and name H4' and segid 2) (residue 7 and name H6 and segid 2) 2.3 1.5 1.5
assign (residue 7 and name H5' and segid 2) (residue 7 and name H6 and segid 2) 1.8 0.7 0.7
assign (residue 7 and name H7* and segid 2) (residue 6 and name H6 and segid 2) 3.7 2.1 2.1

assign (residue 7 and name H7* and segid 2) (residue 7 and name H1' and segid 2) 3.3 1.5 1.5
assign (residue 7 and name H7* and segid 2) (residue 7 and name H6 and segid 2) 2.7 1.1 1.1
assign (residue 8 and name H1' and segid 2) (residue 8 and name H6 and segid 2) 2.6 1.0 1.0
assign (residue 8 and name H2' and segid 2) (residue 8 and name H6 and segid 2) 3.0 1.2 1.2
assign (residue 8 and name H2' and segid 2) (residue 8 and name H1' and segid 2) 2.1 0.9 0.9
assign (residue 8 and name H2" and segid 2) (residue 8 and name H1' and segid 2) 2.5 1.0 1.0
assign (residue 8 and name H2" and segid 2) (residue 8 and name H6 and segid 2) 2.4 1.5 1.5
assign (residue 8 and name H4' and segid 2) (residue 8 and name H1' and segid 2) 3.2 1.3 1.3
assign (residue 8 and name H5 and segid 2) (residue 8 and name H6 and segid 2) 2.4 1.0 1.0
assign (residue 8 and name H5' and segid 2) (residue 8 and name H1' and segid 2) 3.2 2.0 2.0
assign (residue 9 and name H2' and segid 2) (residue 9 and name H6 and segid 2) 3.6 1.4 1.4
assign (residue 9 and name H2" and segid 2) (residue 9 and name H1' and segid 2) 3.1 1.3 1.3
assign (residue 9 and name H2" and segid 2) (residue 9 and name H3' and segid 2) 2.5 1.0 1.0
assign (residue 9 and name H2" and segid 2) (residue 9 and name H6 and segid 2) 3.4 1.4 1.4
assign (residue 9 and name H3' and segid 2) (residue 9 and name H1' and segid 2) 2.3 1.5 1.5
assign (residue 9 and name H3' and segid 2) (residue 9 and name H6 and segid 2) 2.6 1.5 1.5
assign (residue 9 and name H3' and segid 2) (residue 10 and name H6 and segid 2) 3.1 2.0 2.0
assign (residue 9 and name H4' and segid 2) (residue 9 and name H6 and segid 2) 2.7 1.1 1.1
assign (residue 9 and name H5' and segid 2) (residue 9 and name H6 and segid 2) 2.7 1.1 1.1
assign (residue 9 and name H7* and segid 2) (residue 8 and name H5 and segid 2) 3.4 2.0 2.0
assign (residue 9 and name H7* and segid 2) (residue 8 and name H6 and segid 2) 3.0 2.0 2.0
assign (residue 9 and name H7* and segid 2) (residue 9 and name H6 and segid 2) 2.5 1.0 1.0
assign (residue 10 and name H2' and segid 2) (residue 10 and name H1' and segid 2) 2.2 0.9 0.9
assign (residue 10 and name H2' and segid 2) (residue 10 and name H6 and segid 2) 2.5 1.0 1.0
assign (residue 10 and name H2" and segid 2) (residue 10 and name H1' and segid 2) 1.8 0.7 0.7
assign (residue 10 and name H2" and segid 2) (residue 10 and name H6 and segid 2) 2.4 1.3 1.3
assign (residue 10 and name H5' and segid 2) (residue 10 and name H6 and segid 2) 2.2 0.9 0.9
assign (residue 10 and name H7* and segid 2) (residue 9 and name H6 and segid 2) 3.9 2.0 2.0
assign (residue 10 and name H7* and segid 2) (residue 10 and name H6 and segid 2) 2.6 1.0 1.0
assign (residue 11 and name H1' and segid 2) (residue 11 and name H6 and segid 2) 2.7 1.1 1.1
assign (residue 11 and name H2' and segid 2) (residue 11 and name H6 and segid 2) 2.3 1.0 1.0
assign (residue 11 and name H2" and segid 2) (residue 11 and name H6 and segid 2) 2.4 1.0 1.0
assign (residue 11 and name H2" and segid 2) (residue 12 and name H6 and segid 2) 2.2 0.9 0.9
assign (residue 11 and name H3' and segid 2) (residue 11 and name H1' and segid 2) 2.8 1.5 1.5
assign (residue 11 and name H3' and segid 2) (residue 11 and name H6 and segid 2) 2.6 1.1 1.1
assign (residue 11 and name H4' and segid 2) (residue 11 and name H6 and segid 2) 3.1 1.2 1.2
assign (residue 11 and name H5 and segid 2) (residue 10 and name H6 and segid 2) 3.2 2.0 2.0
assign (residue 11 and name H5 and segid 2) (residue 11 and name H6 and segid 2) 2.3 0.9 0.9
assign (residue 11 and name H5' and segid 2) (residue 11 and name H6 and segid 2) 2.0 0.8 0.8
assign (residue 12 and name H1' and segid 2) (residue 12 and name H6 and segid 2) 2.7 1.1 1.1
assign (residue 12 and name H2' and segid 2) (residue 12 and name H1' and segid 2) 2.4 0.9 0.9
assign (residue 12 and name H2' and segid 2) (residue 12 and name H2" and segid 2) 2.0 0.8 0.8
assign (residue 11 and name H2' and segid 2) (residue 12 and name H3' and segid 2) 2.0 0.8 0.8
assign (residue 12 and name H2" and segid 2) (residue 12 and name H1' and segid 2) 2.1 0.8 0.8
assign (residue 12 and name H2" and segid 2) (residue 12 and name H3' and segid 2) 2.4 1.0 1.0
assign (residue 11 and name H2" and segid 2) (residue 12 and name H6 and segid 2) 2.4 1.0 1.0
assign (residue 12 and name H4' and segid 2) (residue 12 and name H1' and segid 2) 2.2 1.1 1.1
assign (residue 12 and name H4' and segid 2) (residue 12 and name H6 and segid 2) 2.6 1.3 1.3
assign (residue 12 and name H5 and segid 2) (residue 11 and name H6 and segid 2) 2.3 1.2 1.2
assign (residue 12 and name H5 and segid 2) (residue 12 and name H6 and segid 2) 2.3 0.9 0.9
assign (residue 12 and name H5' and segid 2) (residue 12 and name H1' and segid 2) 2.8 2.0 2.0

Exchangeable proton distance restraints file

assign (residue 2 and name H2 and segid 1) (residue 27 and name H3 and segid 1) 2.5 1.3 1.3
assign (residue 3 and name H1 and segid 1) (residue 24 and name H3 and segid 1) 4.1 2.0 2.0
assign (residue 4 and name H8 and segid 1) (residue 4 and name H3 and segid 2) 3.9 2.0 2.0

assign (residue 4 and name H1 and segid 1) (residue 3 and name H1 and segid 1) 3.9 2.5 2.5
assign (residue 4 and name H1 and segid 1) (residue 21 and name H3 and segid 1) 4.0 1.6 1.6
assign (residue 4 and name H41 and segid 1) (residue 4 and name H3 and segid 2) 3.3 1.8 1.8
assign (residue 4 and name H42 and segid 1) (residue 4 and name H3 and segid 2) 2.1 1.5 1.5
assign (residue 4 and name H1 and segid 1) (residue 5 and name H3 and segid 2) 4.3 3.0 3.0
assign (residue 5 and name H2 and segid 1) (residue 21 and name H3 and segid 1) 3.9 2.0 2.0
assign (residue 5 and name H2' and segid 1) (residue 6 and name H3 and segid 2) 3.9 3.0 3.0
assign (residue 6 and name H2 and segid 1) (residue 20 and name H3 and segid 1) 3.7 1.5 1.5
assign (residue 6 and name H2' and segid 1) (residue 7 and name H3 and segid 2) 3.7 3.0 3.0
assign (residue 6 and name H2" and segid 1) (residue 7 and name H3 and segid 2) 3.8 2.5 2.5
assign (residue 6 and name H8 and segid 1) (residue 20 and name H3 and segid 1) 3.2 3.0 3.0
assign (residue 6 and name H61 and segid 1) (residue 6 and name H3 and segid 2) 3.6 1.4 1.4
assign (residue 6 and name H61 and segid 1) (residue 7 and name H3 and segid 2) 3.3 2.0 2.0
assign (residue 6 and name H62 and segid 1) (residue 7 and name H3 and segid 2) 3.6 2.5 2.5
assign (residue 7 and name H2 and segid 1) (residue 19 and name H3 and segid 1) 2.2 1.5 1.5
assign (residue 7 and name H61 and segid 1) (residue 7 and name H3 and segid 2) 3.3 1.5 1.5
assign (residue 7 and name H62 and segid 1) (residue 7 and name H3 and segid 2) 3.2 1.5 1.5
assign (residue 8 and name H8 and segid 1) (residue 8 and name H3 and segid 2) 2.9 1.5 1.5
assign (residue 8 and name H1 and segid 1) (residue 17 and name H3 and segid 1) 3.2 2.0 2.0
assign (residue 8 and name H1 and segid 1) (residue 19 and name H3 and segid 1) 3.5 2.5 2.5
assign (residue 9 and name H2 and segid 1) (residue 17 and name H3 and segid 2) 2.3 0.9 0.9
assign (residue 9 and name H8 and segid 1) (residue 9 and name H3 and segid 2) 3.9 1.6 1.6
assign (residue 9 and name H61 and segid 1) (residue 9 and name H3 and segid 2) 3.1 1.2 1.2
assign (residue 10 and name H2' and segid 1) (residue 11 and name H3 and segid 2) 3.1 2.5 2.5
assign (residue 10 and name H61 and segid 1) (residue 10 and name H3 and segid 2) 3.4 1.4 1.4
assign (residue 15 and name H3 and segid 1) (residue 16 and name H3 and segid 1) 3.8 1.5 1.5
assign (residue 15 and name H7* and segid 1) (residue 11 and name H1 and segid 1) 3.7 1.5 1.5
assign (residue 15 and name H7* and segid 1) (residue 15 and name H3 and segid 1) 3.9 1.6 1.6
assign (residue 16 and name H7* and segid 1) (residue 11 and name H1 and segid 1) 3.8 1.5 1.5
assign (residue 16 and name H7* and segid 1) (residue 16 and name H3 and segid 1) 3.7 1.5 1.5
assign (residue 17 and name H7* and segid 1) (residue 8 and name H1 and segid 1) 4.2 1.7 1.7
assign (residue 19 and name H7* and segid 1) (residue 19 and name H3 and segid 1) 2.7 1.1 1.1
assign (residue 20 and name H3 and segid 1) (residue 19 and name H3 and segid 1) 3.7 2.5 2.5
assign (residue 21 and name H7* and segid 1) (residue 21 and name H3 and segid 1) 3.2 1.8 1.8
assign (residue 22 and name H41 and segid 1) (residue 4 and name H1 and segid 1) 3.0 1.5 1.5
assign (residue 23 and name H41 and segid 1) (residue 3 and name H1 and segid 1) 3.3 1.8 1.8
assign (residue 2 and name H6 and segid 2) (residue 2 and name H3 and segid 2) 4.5 3.0 3.0
assign (residue 2 and name H7* and segid 2) (residue 2 and name H3 and segid 2) 3.1 1.5 1.5
assign (residue 3 and name H3 and segid 2) (residue 3 and name H8 and segid 1) 3.9 2.0 2.0
assign (residue 5 and name H3 and segid 2) (residue 4 and name H3 and segid 2) 3.9 3.0 3.0
assign (residue 6 and name H6 and segid 2) (residue 6 and name H3 and segid 2) 3.6 2.0 2.0
assign (residue 6 and name H7* and segid 2) (residue 6 and name H3 and segid 2) 3.7 2.0 2.0
assign (residue 7 and name H3 and segid 2) (residue 6 and name H3 and segid 2) 3.6 2.0 2.0
assign (residue 7 and name H7* and segid 2) (residue 7 and name H3 and segid 2) 2.8 1.5 1.5
assign (residue 7 and name H3 and segid 2) (residue 8 and name H3 and segid 2) 3.2 2.5 2.5
assign (residue 9 and name H3 and segid 2) (residue 8 and name H3 and segid 2) 3.5 2.0 2.0
assign (residue 10 and name H7* and segid 2) (residue 10 and name H3 and segid 2) 4.0 2.5 2.5
assign (residue 10 and name H7* and segid 2) (residue 11 and name H3 and segid 2) 3.6 3.0 3.0

A4. Modifications to CNS files

Topology file

```
MASS  SH   32.06! sulfur
MASS  NS   14.00670! nitrogen in ring >N-
MASS  HS    1.00800! sulfur

! PCYT
MASS  N1C   14.00670! nitrogen in ring >N-
MASS  C2C   12.01100! (prev CN)
MASS  C4C   12.01100! (prev CA)
MASS  C5C   12.011! (prev CF)
MASS  C6C   12.011! (prev CF)
MASS  N4C   14.00670! nitrogen in -NH2
! STHY
MASS  N1T   14.00670! nitrogen in ring >N-
MASS  N3T   14.00670! nitrogen in ring >N-
MASS  C2T   12.01100! (prev CN)
MASS  C4T   12.01100! (prev CN)
MASS  C5T   12.011! (prev CS)
MASS  C6T   12.011! (prev CF)
MASS  CC3E  12.01100! (prev CF)

RESIdue PCYT
 GROUp
 ATOM P   TYPE=P    CHARGE=1.20   END
 ATOM O1P TYPE=O1P  CHARGE=-0.391  END
 ATOM O2P TYPE=O2P  CHARGE=-0.40  END
 ATOM O5' TYPE=O5R  CHARGE=-0.659  END
 GROUp
 ATOM C5' TYPE=C5R  CHARGE=0.213 END
 ATOM H5' TYPE=H    CHARGE=0.044  END
 ATOM H5" TYPE=H    CHARGE=0.044  END
 GROUp
 ATOM C4' TYPE=C4R  CHARGE=0.095  END
 ATOM H4' TYPE=H    CHARGE=0.082  END
 ATOM O4' TYPE=O4R  CHARGE=-0.30  END
 ATOM C1' TYPE=C1R  CHARGE=0.255  END
 ATOM H1' TYPE=H    CHARGE=0.188  END
 GROUp
 ATOM N1   TYPE=N1C  CHARGE=-0.711 EXCLUSION=( C4 ) END
 ATOM C6   TYPE=C6C  CHARGE=0.924 EXCLUSION=( N3 ) END
 ATOM H6   TYPE=H    CHARGE=0.192 END
 GROUp
 ATOM C2   TYPE=C2C  CHARGE=1.065   EXCLUSION=( C5 ) END
 ATOM O2   TYPE=ON   CHARGE=-0.494  END
 GROUp
 ATOM N3   TYPE=NC   CHARGE=-1.034  END
 ATOM H3   TYPE=HN   CHARGE=0.507   END
 ATOM C4   TYPE=C4C  CHARGE=1.309   END
 GROUp
 ATOM N4   TYPE=N4C  CHARGE=-1.257  END
 ATOM H41  TYPE=H2   CHARGE=0.577   END
 ATOM H42  TYPE=H2   CHARGE=0.618   END
 GROUp
 ATOM C5   TYPE=C5C  CHARGE=-1.181  END !CHRG
```

```
 ATOM H5   TYPE=H    CHARGE=0.339   END
GROUp
 ATOM C2' TYPE=C2R  CHARGE=-0.308   END
 ATOM H2' TYPE=H    CHARGE=0.115   END
 ATOM O2' TYPE=O2R  CHARGE=-0.40   END
 ATOM HO2' TYPE=HO  CHARGE=0.25    END
GROUp
 ATOM C3' TYPE=C3R  CHARGE=0.130  END
 ATOM H3' TYPE=H    CHARGE=0.143   END
GROUp
 ATOM O3' TYPE=O3R  CHARGE=-0.641   END
```

```
BOND P   O1P        BOND P   O2P        BOND P    O5'
BOND O5' C5'        BOND C5' C4'        BOND C4'  O4'
BOND C4' C3'        BOND O4' C1'        BOND C1'  N1
BOND C1' C2'        BOND N1  C2         BOND N1   C6
BOND N3  H3         BOND C2  N3         BOND N3   C4
BOND C4  N4         BOND N4  H41        BOND N4   H42
BOND C2  O2         BOND C4  C5         BOND C5   C6
BOND C2' C3'        BOND C3' O3'        BOND C2'  O2'
BOND C6  H6         BOND C5  H5         BOND O2'  HO2'
BOND C5' H5'        BOND C5' H5"        BOND C4'  H4'
BOND C3' H3'        BOND C2' H2'        BOND C1'  H1'
```

```
DIHEdral P   O5' C5' C4'       DIHEdral O5' C5' C4' O4'
DIHEdral O5' C5' C4' C3'       DIHEdral C3' C4' O4' C1'
DIHEdral C4' O4' C1' C2'       DIHEdral O4' C1' C2' C3'
DIHEdral C1' C2' C3' C4'       DIHEdral O4' C4' C3' O3'
DIHEdral C5' C4' C3' C2'       DIHEdral O2' C2' C3' O3'
DIHEdral O4' C1' N1  C2        DIHEdral C3' C2' O2' H2'
DIHEdral  C5' C4' C3' O3'      DIHEdral  C4' O4' C1' N1
```

```
IMPRoper C5  C4  N4  H41        IMPRoper C1' C2  C6  N1
IMPRoper O2  N1  N3  C2         IMPRoper N4  N3  C5  C4
IMPRoper N1  C2  N3  C4         IMPRoper C2  N3  C4  C5
IMPRoper N3  C4  C5  C6         IMPRoper C4  C5  C6  N1
IMPRoper C5  C6  N1  C2         IMPRoper C6  N1  C2  N3
IMPRoper H42 C4  H41 N4         IMPRoper H3  C2  C4  N3
IMPRoper H5  C4  C6  C5         IMPRoper H6  N1  C5  C6
```

```
!RIBOSE IMPROPERS
IMPRoper   H1' C2' O4' N1 !C1'
IMPRoper   H2' C3' C1' O2' !C2'
IMPRoper   H3' C4' C2' O3' !C3'
IMPRoper   H4' C5' C3' O4' !C4'
IMPRoper   H5' O5' H5" C4' !C5'
```

```
 END {PCYT}
 !------------------------------------------------------------

RESIdue STHY
 GROUp
 ATOM P   TYPE=P    CHARGE=1.20   END
 ATOM O1P TYPE=O1P  CHARGE=-0.40   END
 ATOM O2P TYPE=O2P  CHARGE=-0.40   END
 ATOM O5' TYPE=O5R  CHARGE=-0.36   END
 GROUp
 ATOM C5' TYPE=C5R  CHARGE=-0.070 END
 ATOM H5' TYPE=H    CHARGE=0.035 END
 ATOM H5" TYPE=H    CHARGE=0.035 END
```

```
GROUp
 ATOM C4' TYPE=C4R  CHARGE=0.065  END
 ATOM H4' TYPE=H    CHARGE=0.035  END
 ATOM O4' TYPE=O4R  CHARGE=-0.30 END
 ATOM C1' TYPE=C1R  CHARGE=0.20   END
 ATOM H1' TYPE=H    CHARGE=0.165  END
GROUp
 ATOM N1   TYPE=N1T  CHARGE=-0.19 EXCLUSION=( C4 ) END
 ATOM C6   TYPE=C6T  CHARGE=0.155 EXCLUSION=( N3 )  END
 ATOM H6   TYPE=H    CHARGE=0.035 END
 GROUp
 ATOM C2   TYPE=C2T  CHARGE=0.35  EXCLUSION=( C5 )  END
 ATOM O2   TYPE=ON   CHARGE=-0.35 END
 GROUp
 ATOM N3   TYPE=N3T  CHARGE=-0.26 END
 ATOM H3   TYPE=HN   CHARGE=0.26  END
 GROUp
 ATOM C4   TYPE=C4T  CHARGE=0.30  END
 ATOM O4   TYPE=ON   CHARGE=-0.30 END
 GROUp
 ATOM C5   TYPE=C5T  CHARGE=-0.035 END
 ATOM C7   TYPE=CC3E CHARGE=-0.070 END
 ATOM H71  TYPE=H   CHARGE=0.035 END
 ATOM H72  TYPE=H   CHARGE=0.035
 ATOM H73  TYPE=H   CHARGE=0.035 END
 GROUP
 ATOM C2' TYPE=C2R  CHARGE=0.115   END
 ATOM H2' TYPE=H    CHARGE=0.035   END
 ATOM O2' TYPE=O2R  CHARGE=-0.40   END
 ATOM HO2' TYPE=HO   CHARGE=0.25    END
 GROUP
 ATOM C3' TYPE=C3S  CHARGE=-0.035  END
 ATOM H3' TYPE=H    CHARGE=0.035   END
 GROUP
 ATOM S3' TYPE=S3R  CHARGE=-0.36  END


BOND P   O1P         BOND P   O2P         BOND P   O5'
BOND O5'  C5'        BOND C5'  C4'        BOND C4'  O4'
BOND C4'  C3'        BOND O4'  C1'        BOND C1'  N1
BOND C1'  C2'        BOND N1  C2          BOND N1  C6
BOND C2  O2          BOND C2  N3          BOND N3   H3
BOND N3  C4          BOND C4  O4          BOND C4  C5
BOND C5  C7          BOND C5  C6          BOND C2'  C3'
BOND C3'  S3'        BOND C2'  O2'        BOND O2'  HO2'
BOND C5'  H5'        BOND C5'  H5"        BOND C3'  H3'
BOND C2'  H2'        BOND C1'  H1'        BOND C4'  H4'
BOND C7  H71          BOND C7  H72         BOND C7  H73
BOND C6  H6


DIHEdral P   O5'  C5'  C4'       DIHEdral O5'  C5'  C4'  O4'
DIHEdral O5'  C5'  C4'  C3'       DIHEdral C3'  C4'  O4'  C1'
DIHEdral C4'  O4'  C1'  C2'       DIHEdral O4'  C1'  C2'  C3'
DIHEdral C1'  C2'  C3'  C4'       DIHEdral O4'  C4'  C3'  S3'
DIHEdral C5'  C4'  C3'  C2'       DIHEdral O2'  C2'  C3'  S3'
DIHEdral O4'  C1'  N1   C2        DIHEdral C3'  C2'  O2'  H2'
DIHEdral  C5'  C4'  C3'  S3'       DIHEdral  C4'  O4'  C1'  N1


IMPRoper O4  N3  C5  C4         IMPRoper C1'  C2  C6  N1
IMPRoper O2  N1  N3  C2         IMPRoper C4  C5  C6  N1
IMPRoper N1  C2  N3  C4         IMPRoper C2  N3  C4  C5
```

```
IMPRoper N3  C4  C5  C6          IMPRoper C5  C6  N1  C2
IMPRoper C6  N1  C2  N3          IMPRoper H3  C2  C4  N3
IMPRoper C7  C4  C6  C5          IMPRoper  H6  N1  C5  C6

!RIBOSE IMPROPERS
IMPRoper   H1'  C2'  O4'  N1 ! C1'
IMPRoper   H2'  C3'  C1'  O2' !C2'
IMPRoper   H3'  C4'  C2'  S3' !C3'
IMPRoper   H4'  C5'  C3'  O4' !C4'
IMPRoper   H5'  O5'  H5" C4' !C5'

END {STHY}
!----------------------------------------------------------------

PRESidue NUC

 MODIFY ATOM -S3' END  !
 ADD BOND -S3' +P
 ADD ANGLE -C3' -S3' +P
 ADD ANGLE -S3' +P  +O1P
 ADD ANGLE -S3' +P  +O2P
 ADD ANGLE -S3' +P  +O5'

END {NUC}
!----------------------------------------------------------------
```

Parameter file

```
BOND C3S  SH      876.00    1.4300
BOND P   S3R    2326.889   1.607
BOND C3S  C2R    2769.190   1.525    ! 0.011 S
BOND S3R  C3D    1982.674   1.431    ! 0.013 S
BOND S3R  C3S    1982.674   1.431
BOND C3S  H      $kchbond  1.09

ANGLe    HO    S3R    P        139.500 107.300
ANGLe    OH    P     S3R       144.300 102.600
ANGLe    C4D   C3S    SH       139.500 111.000
ANGLe    C4D   C3D   S3R        139.500 111.000
ANGLe    C4R   C3S    SH       139.500 111.000
ANGLe    C2D   C3S    SH       139.500 111.000
ANGLe    C2R   C3S    SH       139.500 111.000
ANGLe    C3S   SH    HO        139.500 107.300
ANGLe    C3S   SH    HO        139.500 107.300
ANGLe    S3R   P     O5R       833.356 104.000
ANGLe    O2P   P     S3R       293.791 108.300
ANGLe    O1P   P     S3R       293.791 107.400
ANGLe    S3R   P     O5R       833.356 104.000
ANGLe    P    S3R   C3S       2089.178 119.700
ANGLe    O4R   C4R   C3S        561.212 105.500
ANGLe    C5R   C4R   C3S        488.878 115.500
ANGLe    C4R   C3S   C2R       1099.976 102.700
ANGLe    C3S   C2R   C1R       1357.996 101.500
ANGLe    C3S   C2R   O2R        357.719 113.300
ANGLe    C4R   C3S   S3R        445.032 110.500
ANGLe    C2R   C3S   S3R        383.726 111.00
ANGLe    O4D   C4D   C3S       1099.976 105.600
```

```
ANGLe    C5D   C4D   C3S      488.878 114.700
ANGLe    C4D   C3S   C2D     1099.976 103.200
ANGLe    C3S   C2D   C1D      650.874 102.700
ANGLe    C4D   C3S   S3R      621.574 110.300 !2.2 S scale from phos.
ANGLe    C4D   C3S   S3R      621.574 110.300 !2.2 S scale from phos.
ANGLe    C2D   C3D   S3R      412.677 110.600 !2.7 S scale from phos.


ANGLe    H     C4R   C3S      $kchangle   107.13
ANGLe    H     C3S   C4R      $kchangle   111.35
ANGLe    H     C3S   S3R      $kchangle   105.87
ANGLe    H     C3S   SH       $kchangle   105.87
ANGLe    H     C3S   C2R      $kchangle   112.27
ANGLe    H     C2R   C3S      $kchangle   111.41
ANGLe    S3R   C3D   H        $kchangle   109.70 !
ANGLe    H     C4D   C3S      $kchangle   106.91
ANGLe    H     C3S   C4D      $kchangle   111.16
ANGLe    H     C3S   S3R      $kchangle   109.34
ANGLe    H     C3S   SH       $kchangle   109.34
ANGLe    H     C3S   C2D      $kchangle   111.98
ANGLe    H     C2D   C3S      $kchangle   111.36
ANGLe    C2C   NC    HN       105.000 116.500 !from param11.dna, 3*keq
ANGLe    C4C   NC    HN       105.000 116.500


DIHEdral  X   C3S   SH   X     1.50 3   0.000


IMPRoper  H   C3S   C1R   O2R   $kchimpr   0   65.000!C2R
IMPRoper  H   C4R   C2R   S3R   $kchimpr   0   60.300!C3R
IMPRoper  H   C5R   C3S   O4R   $kchimpr   0   70.300!C4R
IMPRoper  H   C3S   H     C1D   $kchimpr   0  -73.500!C2D
IMPRoper  H   C4D   C2D   S3R   $kchimpr   0   62.660!C3D
IMPRoper  H   C5D   C3S   O4D   $kchimpr   0   70.220!C4D


DIHEdral  O3S  P    O5R   C5R   1.41 3   24    ! alpha !P (20.3)
DIHEdral  O5R  C5R  C4R   C3S   12.24 3  18    ! gamma !S (6.9)
DIHEdral  C4R  C3R  O3S   P     7.88 0 -153    ! eps  !P (8.6)
DIHEdral  C3S  C3R  P     O5R   1.75 3   33    ! zeta !P (18.3)
DIHEdral  S3R  P    O5R   C5D   1.41 3   6.0 !DNA
DIHEdral  O5R  C5D  C4D   C3S   12.42 3  18.3
DIHEdral  C4D  C3S  S3R   P     7.88 0 214.0
DIHEdral  C3S  S3R  P     O5R   1.75 3   0.3
DIHEdral  C5R  C4R  C3S   S3R   30.12 0  81.1  ! delta ! c3'-endo S (4.4)
DIHEdral  O4R  C4R  C3S   S3R   33.10 0 201.8  ! 4.2 ! c3'-endo S
DIHEdral  O4R  C1R  C2R   C3S   24.28 0 335.4  ! 4.9 ! c3'-endo S
DIHEdral  C1R  C2R  C3S   C4R   74.36 0  35.9  ! 2.8 ! c3'-endo S
DIHEdral  C2R  C3S  C4R   O4R   60.67 0 324.7  ! 3.1 ! c3'-endo S
DIHEdral  C3S  C4R  O4R   C1R   22.42 0  20.5  ! 5.1 ! c3'-endo S
DIHEdral  C5R  C4R  C3S   C2R   60.67 0 204.0  ! 3.1 ! c3'-endo S
DIHEdral  S3R  C3S  C2R   O2R   28.79 0  44.3  ! 4.5 ! c3'-endo S
DIHEdral  C5D  C4D  C3S   S3R   36.44 0 145.2  ! delta ! c2'-endo S (4.0)
DIHEdral  O4D  C1D  C2D   C3S   24.28 0  32.8  ! 4.9 ! c2'-endo S
DIHEdral  O4D  C4D  C3S   S3R   31.53 0 265.8  ! 4.3 ! c2'-endo S
DIHEdral  C1D  C2D  C3S   C4D   44.99 0 326.9  ! 3.6 ! c2'-endo S
DIHEdral  C5D  C4D  C3S   C2D   34.68 0 262.0  ! 4.1 ! c2'-endo S


IMPRoper  O3S  X    X     C3S   94.5 0  35.260
IMPRoper  C5R  O4R  C3S   C4R   94.5 0  35.260


NONBonded C3S    0.0900 3.2970    0.0900 3.2970
NONBonded HS     0.0045 2.6160    0.0045 2.6160
NONBonded S3R     0.2304 2.7290    0.2304 2.7290
```

NONBonded NS   0.1600 2.8591   0.1600 2.8591
NONBonded SD   0.3515 2.6727   0.3515 2.6727 !

nbfix HS S3R   0.05   0.1   0.05   0.1
nbfix H   S3R   0.05   0.1   0.05   0.1

Link File

link nuc head - sthy tail + * end
link nuc head - pcyt tail + * end

first 5ter tail + sthy end
first 5ter tail + pcyt end

last 3ter head - sthy end
last 3ter head - pcyt end

References

References

1. W. Saenger, *Principles of Nucleic Acid Structure*. (1984), New York, USA: Springer Verlag.

2. O. T. Avery, C. M. MacLeod, and M. McCarty, Studies on the chemical nature of the substance inducing transformation of pneumococcal types. Induction of transformation by a desoxyribonucleic acid fraction isolated from pneumococcus Type III. *Journal of Experimental Medicine*, (1944) **79**: 137-158.

3. J. D. Watson and F. H. Crick, The structure of DNA. *Cold Spring Harbor Symposia on Quantitative Biology*, (1953) **18**: 123-131.

4. R. E. Franklin and R. G. Gosling, Evidence for 2-chain helix in crystalline structure of sodium deoxyribonucleate. *Nature*, (1953) **172**(4369): 156-157.

5. V. R. Cooper, T. Thonhauser, A. Puzder, E. Schroder, B. I. Lundqvist, and D. C. Langreth, Stacking Interactions and the Twist of DNA. *Journal of the American Chemical Society*, (2007) **130**(4): 1304-1308.

6. E. T. Kool, Hydrogen bonding, base stacking and steric effects in DNA replication. *Annual Review of Biophysics and Biomolecular Structure*, (2001) **30**(1): 1-22.

7. C. P. Fenna, V. J. Wilkinson, J. R. P. Arnold, R. Cosstick, and J. Fisher, The effect of 2'-fluorine substitutions on DNA i-motif conformation and stability. *Chemical Communications*, (2008)(30): 3567-3569.

8. G. M. Blackburn and M. J. Gait, eds. *Nucleic Acids in Chemistry and Biology*. Third Edition ed. (2006), RSC Publishing: Cambridge.

9. S. Neidle, *Nucleic Acid Structure*. (1999), Oxford: Oxford University Press.

10. G. Felsenfeld, D. R. Davies, and A. Rich, Formation of a three-stranded polynucleotide molecule. *Journal of the American Chemical Society*, (1957) **79**: 2023-2024.

11. K. Hoogsteen, The structure of crystals containing a hydrogen-bonded complex of 1-methylthymine and 9-methyladenine. *Acta Crystallographica*, (1959) **12**(10): 822-823.

12. M. N. Lipsett, Interactions of poly C and guanine trinucleotide. *Biochemical and Biophysical Research Communications*, (1963) **11**(3): 224-228.

13. M. N. Lipsett, Complex formation between polycytidylic acid and guanine oligonucleotides. *Journal of Biological Chemistry*, (1964) **239**(4): 1256-1260.

14. M. J. Chamberlain and D. L. Patterson, Physical and chemical characterization of ordered complexes formed between polyinosinic acid, polycytidylic acid and their deoxyribo-analogues. *Journal of Molecular Biology*, (1965) **12**(2): 410-428.

15. A. Rich, A hybrid helix containing both deoxyribose and ribose polynucleotides and its relation to the transfer of information between the nucleic acids. *Proceedings of the National Academy of Sciences of the United States of America*, (1960) **46**(8): 1044-1053.

16. A. R. Morgan and R. D. Wells, Specificity of 3-stranded complex formation between double-stranded DNA and single-stranded RNA containing repeating nucleotide sequences. *Journal of Molecular Biology*, (1968) **37**(1): 63-80.

17. S. Arnott and E. Selsing, Structures for the polynucleotide complexes poly(dA):poly(dT) and poly(dT):poly(dA):poly(dT). *Journal of Molecular Biology*, (1974) **88**(2): 509-521.

18. A. Jain, G. Wang, and K. M. Vasquez, DNA triple helices: Biological consequences and therapeutic potential. *Biochimie*, (2008) **90**(8): 1117-1130.

19. J. Bijapur, M. D. Keppler, S. Bergqvist, T. Brown, and K. R. Fox, 5-(1-propargylamino)-2 '-deoxyuridine (U-P): a novel thymidine analogue for generating DNA triplexes with increased stability. *Nucleic Acids Research*, (1999) **27**(8): 1802-1809.

20. J. Jaumot, A. Avino, R. Eritja, R. Tauler, and R. Gargallo, Resolution of parallel and antiparallel oligonucleotide triple helices formation and melting processes by multivariate curve resolution. *Journal of Biomolecular Structure & Dynamics*, (2003) **21**(2): 267-278.

21. J. S. Sun and C. Helene, Oligonucleotide-directed triple-helix formation. *Current Opinion in Structural Biology*, (1993) **3**(3): 345-356.

22. D. A. Rusling, P. A. Rachwal, T. Brown, and K. R. Fox, The stability of triplex DNA is affected by the stability of the underlying duplex. *Biophysical Chemistry*, (2009) **145**(2-3): 105-110.

23. P. R. Hoyne, A. M. Gacy, C. T. McMurray, and L. J. Maher, III, Stabilities of intrastrand pyrimidine motif DNA and RNA triple helices. *Nucleic Acids Res.*, (2000) **28**(3): 770-775.

24. Z.-J. Tan and S.-J. Chen, *Importance of Diffuse Metal Ion Binding to RNA*. Structural and Catalytic Roles of Metal Ions in RNA. Vol. 9. (2011). 101-124.

25. A. R. Srinivasan and W. K. Olson, Molecular models of nucleic acid triple helixes. I. DNA and RNA backbone complexes. *Journal of the American Chemical Society*, (1998) **120**(3): 484-491.

26. H. Han and P. B. Dervan, Different conformational families of pyrimidine-purine-pyrimidine triple helices depending on backbone composition. *Nucleic Acids Research*, (1994) **22**(14): 2837-2844.

27. I. Radhakrishnan and D. J. Patel, Solution structure of a purine-purine-pyrimidine DNA triplex containing GGC and TAT triples. *Structure*, (1993) **1**(2): 135-152.

28. D. Vlieghe, L. Van Meervelt, A. Dautant, B. Gallois, G. Precigoux, and O. Kennard, Parallel and antiparallel (GGC)$_2$ triple helix fragments in a crystal structure. *Science*, (1996) **273**(5282): 1702-1705.

29. M. D. Frank-Kamenetskii and S. M. Mirkin, Triplex DNA structures. *Annual Review of Biochemistry*, (1995) **64**: 65-95.

30. J. C. Hanvey, M. Shimizu, and R. D. Wells, Intramolecular DNA triplexes in supercoiled plasmids. *Proceedings of the National Academy of Sciences of the United States of America* (1988) **85**(17): 6292-6296.

31. J. Zhao, A. Bacolla, G. Wang, and K. M. Vasquez, Non-B DNA structure-induced genetic instability and evolution. *Cellular and Molecular Life Sciences*, (2010) **67**(1): 43-62.

32. C. H. Gotfredsen, P. Schultze, and J. Feigon, Solution structure of an intramolecular pyrimidine-purine-pyrimidine triplex containing an RNA third strand. *Journal of the American Chemical Society*, (1998) **120**(18): 4281-4289.

33. R. W. Roberts and D. M. Crothers, Stability and properties of double and triple helices - dramatic effects of RNA or DNA backbone composition. *Science*, (1992) **258**(5087): 1463-1466.

34. G. M. Hashem, J.-D. Wen, Q. Do, and D. M. Gray, Evidence from CD spectra and melting temperatures for stable Hoogsteen-paired oligomer duplexes derived from DNA and hybrid triplexes. *Nucleic Acids Research*, (1999) **27**(16): 3371-3379.

35. C. Escude, J. C. Francois, J. S. Sun, G. Ott, M. Sprinzl, T. Garestier, and C. Helene, Stability of triple helices containing RNA and DNA strands - experimental and molecular modeling studies. *Nucleic Acids Research*, (1993) **21**(24): 5547-5553.

36. M. J. J. Blommers, F. Natt, W. Jahnke, and B. Cuenoud, Dual recognition of double-stranded DNA by 2'-aminoethoxy-modified oligonucleotides: the solution structure of an intramolecular triplex obtained by NMR spectroscopy. *Biochemistry*, (1998) **37**(51): 17714-17725.

37. J. L. Asensio, R. Carr, T. Brown, and A. N. Lane, Conformational and thermodynamic properties of parallel intramolecular triple helices containing a DNA, RNA, or 2'-OMe DNA third strand. *Journal of the American Chemical Society*, (1999) **121**(48): 11063-11070.

38. F. Eckstein, *Oligonucleotides and Analogues*. A Practiacal Approach. (1991): Oxford University Press.

39. A. H. El-Sagheer and T. Brown, New strategy for the synthesis of chemically modified RNA constructs exemplified by hairpin and hammerhead ribozymes. *Proceedings of the National Academy of Sciences*, (2010) **107**(35): 15329-15334.

40. T. J. Povsic and P. B. Dervan, Triple helix formation by oligonucleotides on DNA extended to the physiological pH range. *Journal of the American Chemical Society*, (1989) **111**(8): 3059-3061.

41. L. E. Xodo, G. Manzini, F. Quadrifoglio, G. A. Van der Marel, and J. H. Van Boom, Effect of 5-methylcytosine on the stability of triple-stranded DNA - a thermodynamic study. *Nucleic Acids Research*, (1991) **19**(20): 5625-5631.

42. P. J. Bates, C. A. Laughton, T. C. Jenkins, D. C. Capaldi, P. D. Roselt, C. B. Reese, and S. Neidle, Efficient triple helix formation by oligodeoxyribonucleotides containing alpha- or beta-2-amino-5-(2-deoxy-D-ribofuranosyl)pyridine residues. *Nucleic Acids Research*, (1996) **24**(21): 4176-4184.

43. S. A. Cassidy, P. Slickers, J. O. Trent, D. C. Capaldi, P. D. Roselt, C. B. Reese, S. Neidle, and K. R. Fox, Recognition of GC base pairs by triplex forming oligonucleotides containing nucleosides derived from 2-aminopyridine. *Nucleic Acids Research*, (1997) **25**(24): 4891-4898.

44. S. Hildbrand, A. Blaser, S. P. Parel, and C. J. Leumann, 5-substituted 2-aminopyridine C-nucleosides as protonated cytidine equivalents: Increasing efficiency and selectivity in DNA triple-helix formation. *Journal of the American Chemical Society*, (1997) **119**(24): 5499-5511.

45. K. R. Fox and T. Brown, Formation of stable DNA triplexes. *Biochemical Society Transactions*, (2011) **39**: 629-634.

46.   A. Ono, P. O. P. Tso, and L. S. Kan, Triplex formation of an oligonucleotide containing 2'-O-methylpseudoisocytidine with a DNA duplex at neutral pH. *Journal of Organic Chemistry*, (1992) **57**(11): 3225-3230.

47.   G. B. Xiang, R. Bogacki, and L. W. McLaughlin, Use of a pyrimidine nucleoside that functions as a bidentate hydrogen bond donor for the recognition of isolated or contiguous G-C base pairs by oligonucleotide-directed triplex formation. *Nucleic Acids Research*, (1996) **24**(10): 1963-1970.

48.   U. Von Krosigk and S. A. Benner, pH-independent triple-helix formation by an oligonucleotide containing a pyrazine donor-donor acceptor base. *Journal of the American Chemical Society*, (1995) **117**(19): 5361-5362.

49.   M. D. Keppler and K. R. Fox, Relative stability of triplexes containing different numbers of TAT and C$^+$GC triplets. *Nucleic Acids Research*, (1997) **25**(22): 4644-4649.

50.   A. B. Staubli and P. B. Dervan, Sequence specificity of the nonnatural pyrido[2,3-D]pyrimidine nucleoside in triplex-helix formation. *Nucleic Acids Research*, (1994) **22**(13): 2637-2642.

51.   T. S. Rao, R. H. Durland, D. M. Seth, M. A. Myrick, V. Bodepudi, and G. R. Revankar, Incorporation of 2'-Deoxy-6-thioguanosine into G-rich oligodeoxyribonucleotides inhibits G-tetrad formation and facilitates triplex formation. *Biochemistry*, (1995) **34**(3): 765-772.

52.   J. F. Milligan, S. H. Krawczyk, S. Wadwani, and M. D. Matteucci, An antiparallel triple helix motif with oligodeoxynucleotides containing 2'-deoxyguanosine and 7-deaza-2'-deoxyxanthosine. *Nucleic Acids Research*, (1993) **21**(2): 327-333.

53.   H. Torigoe, R. Shimizume, A. Sarai, and H. Shindo, Triplex formation of chemically modified homopyrimidine oligonucleotides: thermodynamic and kinetic studies. *Biochemistry*, (1999) **38**(44): 14653-14659.

54.   S. Karkare and D. Bhatnagar, Promising nucleic acid analogs and mimics: characteristic features and applications of PNA, LNA, and morpholino. *Applied Microbiology and Biotechnology*, (2006) **71**(5): 575-586.

55.   J. J. Sorensen, J. T. Nielsen, and M. Petersen, Solution structure of a dsDNA:LNA triplex. *Nucleic Acids Research*, (2004) **32**(20): 6078-6085.

56.   M. E. Hansen, T. Bentin, and P. E. Nielsen, High-affinity triplex targeting of double stranded DNA using chemically modified peptide nucleic acid oligomers. *Nucleic Acids Research*, (2009) **37**(13): 4498-4507.

57. S. Shakeel, S. Karim, and A. Ali, Peptide nucleic acid (PNA) - a review. *Journal of Chemical Technology and Biotechnology*, (2006) **81**(6): 892-899.

58. L. Lacroix, P. B. Arimondo, M. Takasugi, C. Helene, and J.-L. Mergny, Pyrimidine morpholino oligonucleotides form a stable triple helix in the absence of magnesium ions. *Biochemical and Biophysical Research Communications*, (2000) **270**(2): 363-369.

59. H. Torigoe, K. Kawahashi, and Y. Tamura, Promotion of triplex formation by morpholino modification: thermodynamic and kinetic studies. *Nucleic Acids Symposium Series*, **3**(1): 157-158.

60. F. A. Buske, J. S. Mattick, and T. L. Bailey, Potential in vivo roles of nucleic acid triple-helices. *RNA Biology*, (2011) **8**(3): 427-439.

61. M. Ohno, T. Fukagawa, J. Lee, and T. Ikemura, Triplex-forming DNAs in the human interphase nucleus visualized in situ by polypurine/polypyrimidine DNA probes and antitriplex antibodies. *Chromosoma*, (2002) **111**(3): 201-213.

62. I. Martianov, A. Ramadass, A. S. Barros, N. Chow, and A. Akoulitchev, Repression of the human dihydrofolate reductase gene by a non-coding interfering transcript. *Nature*, (2007) **445**(7128): 666-670.

63. L. J. Collins, C. G. Kurland, P. Biggs, and D. Penny, The modern RNP world of eukaryotes. *Journal of Heredity*, (2009) **100**(5): 597-604.

64. B. M. Lunde, C. Moore, and G. Varani, RNA-binding proteins: modular design for efficient function. *Nat Rev Mol Cell Biol*, (2007) **8**(6): 479-490.

65. K. J. Hampel, G. D. Burkholder, and J. S. Lee, Plasmid dimerization mediated by triplex formation between polypyrimidine polypurine repeats. *Biochemistry*, (1993) **32**(4): 1072-1077.

66. F. X. Barre, S. Ait-Si-Ali, C. Giovannangeli, R. Luis, P. Robin, L. L. Pritchard, C. Helene, and A. Harel-Bellan, Unambiguous demonstration of triple-helix-directed gene modification. *Proceedings of the National Academy of Sciences of the United States of America*, (2000) **97**(7): 3084-3088.

67. M. Duca, P. Vekhoff, K. Oussedik, L. Halby, and P. B. Arimondo, The triple helix: 50 years later, the outcome. *Nucleic Acids Research*, (2008) **36**(16): 5123-5138.

68. S. Obika, Development of bridged nucleic acid analogues for antigene technology. *Chemical and Pharmaceutical Bulletin*, (2004) **52**(12): 1399-1404.

69.   D. Praseuth, A. L. Guieysse, and C. Helene, Triple helix formation and the antigene strategy for sequence-specific control of gene expression. *Biochimica et Biophysica Acta, Gene Structure and Expression*, (1999) **1489**(1): 181-206.

70.   M. Faria and C. Giovannangeli, Triplex-forming molecules: from concepts to applications. *Journal of Gene Medicine*, (2001) **3**(4): 299-310.

71.   P. B. Arimondo, C. J. Thomas, K. Oussedik, B. Baldeyrou, C. Mahieu, L. Halby, D. Guianvarc'h, A. Lansiaux, S. M. Hecht, C. Bailly, and C. Giovannangeli, Exploring the cellular activity of camptothecin-triple-helix-forming oligonucleotide conjugates. *Molecular and Cellular Biology*, (2006) **26**(1): 324-333.

72.   G. M. Carbone, E. M. McGuffie, A. Collier, and C. V. Catapano, Selective inhibition of transcription of the Ets2 gene in prostate cancer cells by a triplex-forming oligonucleotide. *Nucleic Acids Research*, (2003) **31**(3): 833-843.

73.   P. W. Hewett, E. L. Daft, C. A. Laughton, S. Ahmad, A. Ahmed, and J. C. Murray, Selective inhibition of the human Tie-1 promoter with triplex-forming oligonucleotides targeted to Ets binding sites. *Molecular Medicine*, (2006) **12**(1-3): 8-16.

74.   J. G. Hacia, P. B. Dervan, and B. J. Wold, Inhibition of klenow fragment DNA-polymerase on double-helical templates by oligonucleotide-directed triple-helix formation. *Biochemistry*, (1994) **33**(20): 6192-6200.

75.   S. Napoli, U. Negri, F. Arcamone, M. L. Capobianco, G. M. Carbone, and C. V. Catapano, Growth inhibition and apoptosis induced by daunomycin-conjugated triplex-forming oligonucleotides targeting the c-myc gene in prostate cancer cells. *Nucleic Acids Research*, (2006) **34**(2): 734-744.

76.   P. P. Chan and P. M. Glazer, Triplex DNA: fundamentals, advances, and potential applications for gene therapy. *Journal of Molecular Medicine*, (1997) **75**(4): 267-282.

77.   J. Y. Chin, J. Y. Kuan, P. S. Lonkar, D. S. Krause, M. M. Seidman, K. R. Peterson, P. E. Nielsen, R. Kole, and P. M. Glazer, Correction of a splice-site mutation in the beta-globin gene stimulated by triplex-forming peptide nucleic acids. *Proceedings of the National Academy of Sciences of the United States of America*, (2008) **105**(36): 13514-13519.

78.   J. Song, Z. Intody, M. Li, and J. H. Wilson, Activation of gene expression by triplex-directed psoralen crosslinks. *Gene*, (2004) **324**: 183-190.

79. M. K. Ghosh, A. Katyal, R. Chandra, and V. Brahmachari, Targeted activation of transcription in vivo through hairpin-triplex forming oligonucleotide in Saccharomyces cerevisiae. *Molecular and Cellular Biochemistry*, (2005) **278**(1-2): 147-155.

80. C. D. Pesce, F. Bolacchi, B. Bongiovanni, F. Cisotta, M. Capozzi, S. Diviacco, F. Quadrifoglio, R. Mango, G. Novelli, G. Mossa, C. Esposito, D. Ombres, G. Rocchi, and A. Bergamini, Anti-gene peptide nucleic acid targeted to proviral HIV-1 DNA inhibits in vitro HIV-1 replication. *Antiviral Research*, (2005) **66**(1): 13-22.

81. C. Rocher, R. Dalibart, T. Letellier, G. Precigoux, and P. Lestienne, Initiation of DNA replication by DNA polymerases from primers forming a triple helix. *Nucleic Acids Research*, (2001) **29**(16): 3320-3326.

82. A. J. Ziemba, M. W. Reed, K. D. Raney, A. B. Byrd, and S. W. Ebbinghaus, A bis-alkylating triplex forming oligonucleotide inhibits intracellular reporter gene expression and prevents triplex unwinding due to helicase activity. *Biochemistry*, (2003) **42**(17): 5013-5024.

83. P. P. Lestienne, Are there three polynucleotide strands in the catalytic centre of DNA polymerases? *Biochimie*, (2009) **91**(11-12): 1523-1530.

84. M. Matteucci, K. Y. Lin, T. Huang, R. Wagner, D. D. Sternbach, M. Mehrotra, and J. M. Besterman, Sequence-specific targeting of duplex DNA using a camptothecin triple helix forming oligonucleotide conjugate and topoisomerase I. *Journal of the American Chemical Society*, (1997) **119**(29): 6939-6940.

85. P. B. Arimondo, D. Guianvarc'h, C. Bailly, A. Boutorine, J. S. Sun, M. Kuwahara, S. Hecht, C. Giovannangeli, T. Garestier, and C. Helene, Design and optimization of topoisomerase poison conjugates of triple helix-forming oligonucleotides for sequence-specific DNA cleavage. *European Journal of Cancer*, (2002) **38**: 414.

86. P. B. Arimondo, P. Moreau, A. Boutorine, C. Bailly, M. Prudhomme, J. S. Sun, T. Garestier, and C. Helene, Recognition and cleavage of DNA by rebeccamycin- or benzopyridoquinoxaline conjugated of triple helix-forming oligonucleotides. *Bioorganic & Medicinal Chemistry*, (2000) **8**(4): 777-784.

87. P. B. Arimondo, C. Bailly, A. S. Boutorine, P. Moreau, M. Prudhomme, J. S. Sun, T. Garestier, and C. Helene, Triple helix-forming oligonucleotides conjugated to indolocarbazole poisons direct topoisomerase I-mediated DNA cleavage to a specific site. *Bioconjugate Chemistry*, (2001) **12**(4): 501-509.

88. M. L. Capobianco, M. De Champdore, F. Arcamone, A. Garbesi, D. Gulanvarc'h, and P. B. Arimondo, Improved synthesis of daunomycin

conjugates with triplex-forming oligonucleotides. The polypurine tract of HIV-1 as a target. *Bioorganic & Medicinal Chemistry*, (2005) **13**(9): 3209-3218.

89.   H. M. Ji and L. M. Smith, Rapid purification of double-stranded DNA by triple-helix-mediated affinity capture. *Analytical Chemistry*, (1993) **65**(10): 1323-1328.

90.   T. Schluep and C. L. Cooney, Purification of plasmids by triplex affinity interaction. *Nucleic Acids Research*, (1998) **26**(19): 4524-4528.

91.   J. W. Gaynor, J. Bentley, and R. Cosstick, Synthesis of the 3 '-thio-nucleosides and subsequent automated synthesis of oligodeoxynucleotides containing a 3 '-*S*-phosphorothiolate linkage. *Nature Protocols*, (2007) **2**(12): 3122-3135.

92.   H. K. Jayakumar, J. L. Buckingham, J. A. Brazier, N. G. Berry, R. Cosstick, and J. Fisher, NMR studies of the conformational effect of single and double 3'-*S*-phosphorothiolate substitutions within deoxythymidine trinucleotides. *Magn. Reson. Chem.*, (2007) **45**(4): 340-345.

93.   J. A. Brazier, J. Fisher, and R. Cosstick, Stabilization of the DNA i-motif structure by incorporation of 3'-*S*-phosphorothiolate linkages. *Angewandte Chemie, International Edition*, (2006) **45**(1): 114-117.

94.   J. Bentley, J. A. Brazier, J. Fisher, and R. Cosstick, Duplex stability of DNA:DNA and DNA:RNA duplexes containing 3 '-S-phosphorothiolate linkages. *Organic & Biomolecular Chemistry*, (2007) **5**(22): 3698-3702.

95.   A. P. G. Beevers, E. M. Witch, B. C. N. M. Jones, R. Cosstick, J. R. P. Arnold, and J. Fisher, Conformational analysis of 3'-*S*-PO$_3$-linked ribo- and deoxyribodinucleoside monophosphates. *Magnetic Resonance in Chemistry*, (1999) **37**(11): 814-820.

96.   J. Buckingham, G. Sabbagh, J. Brazier, J. Fisher, and R. Cosstick, Control of DNA conformation using 3'-*S*-phosphorothiolate-modified linkages. *Nucleosides Nucleotides & Nucleic Acids*, (2005) **24**(5-7): 491-495.

97.   A. P. G. Beevers, K. J. Fettes, I. A. O'Neil, S. M. Roberts, J. R. P. Arnold, R. Cosstick, and J. Fisher, Probing the effect of a 3 '-*S*-phosphorothiolate link on the conformation of a DNA : RNA hybrid; implications for antisense drug design. *Chemical Communications*, (2002)(14): 1458-1459.

98.   J. S. Vyle, B. A. Connolly, D. Kemp, and R. Cosstick, Sequence- and strand-specific cleavage in oligodeoxyribonucleotides and DNA containing 3'-thiothymidine. *Biochemistry*, (1992) **31**(11): 3012-3018.

99. F. J. M. Van De Ven and C. W. Hilbers, Nucleic acids and nuclear magnetic resonance. *European Journal of Biochemistry*, (1988) **178**(1): 1-38.

100. Meena, M. Sam, K. Pierce, J. W. Szostak, and L. W. McLaughlin, 2',3'-Dideoxy-3'-thionucleoside Triphosphates: Syntheses and Polymerase Substrate Activities. *Organic Letters*, (2007) **9**(6): 1161-1163.

101. S. M. Freier and K. H. Altmann, The ups and downs of nucleic acid duplex stability: structure-stability studies on chemically-modified DNA:RNA duplexes. *Nucleic Acids Research*, (1997) **25**(22): 4429-4443.

102. B. A. Luxon and D. G. Gorenstein, Comparison of X-ray and NMR-determined nucleic acid structures. *Nuclear Magnetic Resonance and Nucleic Acids*, (1995) **261**: 45-73.

103. V. V. Krishnan and B. Rupp, *Macromolecular Structure Determination: Comparison of X-ray Crystallography and NMR Spectroscopy*, in *Macromolecular Structure Determination: Comparison of X-ray Crystallography and NMR Spectroscopy*. (2001), John Wiley & Sons, Ltd.

104. H. DeVoe, The theory of hypochromism of biopolymers: calculated spectra for DNA. *Annals of the New York Academy of Sciences*, (1969) **158**(1): 298-307.

105. J. D. Puglisi and I. Tinoco, Jr., Absorbance melting curves of RNA. *Methods in Enzymology*, (1989) **180**: 304-325.

106. B. Bardiaux, T. Malliavin, M. Nilges, and A. Mazur, Comparison of different torsion angle approaches for NMR structure determination. *Journal of Biomolecular NMR*, (2006) **34**(3): 153-166.

107. L. J. Rinkel and C. Altona, Conformational analysis of the deoxyribofuranose ring in DNA by means of sums of proton-proton coupling constants - a graphical method. *Journal of Biomolecular Structure & Dynamics*, (1987) **4**(4): 621-649.

108. M. Remin and D. Shugar, Conformation of exocyclic 5'-CH_2OH in nucleosides and nucleotides in aqueous solution from specific assignments of H5' and H5" signals in the NMR spectra. *Biochemical and Biophysical Research Communications*, (1972) **48**(3): 636-&.

109. S. S. Wijmenga and B. N. M. van Buuren, The use of NMR methods for conformational studies of nucleic acids. *Progress in Nuclear Magnetic Resonance Spectroscopy*, (1998) **32**(4): 287-387.

110. K. Wuthrich, *NMR of Proteins and Nucleic Acids*. (1986), New York, USA: John Wiley & Sons.

111. G. Varani, F. Aboul-ela, and F. H. T. Allain, NMR investigation of RNA structure. *Progress in Nuclear Magnetic Resonance Spectroscopy*, (1996) **29**(1-2): 51-127.

112. J. Zhang, A. Spring, and M. W. Germann, Facilitated assignment of adenine H2 resonances in oligonucleotides using homonuclear long-range couplings. *Journal of the American Chemical Society*, (2009) **131**(15): 5380-5381.

113. B. Gopalakrishnan and M. Bansals, Comparison of interproton distances in DNA models with nuclear Overhauser enhancement data. *Journal of Biosciences*, (1985) **8**(3-4): 603-614.

114. M. R. Conte, T. C. Jenkins, and A. N. Lane, Interaction of minor-groove-binding diamidine ligands with an asymmetric DNA duplex. *European Journal of Biochemistry*, (1995) **229**(2): 433-444.

115. C. Glemarec, A. Nyilas, C. Sund, and J. Chattopadhyaya, A comparative conformational study of thymidylyl(3'----5')-thymidine, thymidylyl(3'----5')-5'-thio-5'- deoxythymidine and thymidinylacetamido-[3'(O)----5'(C)]-5'-deoxythymidine. *Journal of Biochemical and Biophysical Methods*, (1990) **21**(4): 311-332.

116. H. P. Spielmann, T. J. Dwyer, J. E. Hearst, and D. E. Wemmer, Solution structures of psoralen monoadducted and cross-linked DNA oligomers by NMR spectroscopy and restrained molecular dynamics. *Biochemistry*, (1995) **34**(40): 12937-12953.

117. N. Ulyanov, W. Bauer, and T. James, High-resolution NMR structure of an AT-rich DNA sequence. *Journal of Biomolecular NMR*, (2002) **22**(3): 265-280.

118. A. T. Brunger, P. D. Adams, G. M. Clore, W. L. DeLano, P. Gros, R. W. Grosse-Kunstleve, J. S. Jiang, J. Kuszewski, M. Nilges, N. S. Pannu, R. J. Read, L. M. Rice, T. Simonson, and G. L. Warren, Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallographica Section D: Biological Crystallography*, (1998) **54**(Pt 5): 905-921.

119. T. L. James, *NMR Determination of Oligonucleotide Structure*, in *Current Protocols in Nucleic Acid Chemistry*. (2001), John Wiley & Sons, Inc.: New York, USA.

120. H. Liu, H. P. Spielmann, N. B. Ulyanov, D. E. Wemmer, and T. L. James, Interproton distance bounds from 2D NOE intensities: effect of experimental noise and peak integration errors. *Journal of Biomolecular NMR*, (1995) **6**(4): 390-402.

121. C. B. Post, R. P. Meadows, and D. G. Gorenstein, On the evaluation of interproton distances for three-dimensional structure determination by NMR using a relaxation rate matrix analysis. *Journal of the American Chemical Society*, (1990) **112**(19): 6796-6803.

122. U. Schmitz, T. L. James, and L. J. Thomas, How to generate accurate solution structures of double-helical nucleic acid fragments using nuclear magnetic resonance and restrained molecular dynamics. *Methods in Enzymology*, (1995) **261**: 3-44.

123. T. L. James, Relaxation matrix analysis of two-dimensional nuclear Overhauser effect spectra. *Current Opinion in Structural Biology*, (1991) **1**(6): 1042-1053.

124. K. Hyz, W. Bocian, R. Kawecki, E. Bednarek, J. Sitkowski, and L. Kozerski, A dumbbell double nicked duplex dodecamer DNA with a PEG6 tether. *Organic & Biomolecular Chemistry*, (2011) **9**(12): 4481-4486.

125. D. M. Heller and A. Giorgetti, NMR Constraints Analyser: a web-server for the graphical analysis of NMR experimental constraints. *Nucleic Acids Research*, (2010) **38**(Web Server): 628-632.

126. H. D. Holtje, W. Sippl, D. Rognan, and G. Folkers, *Molecular Modelling; Basic Principles and Applications*. (2003), Weinheim, Germany: Wiley - VHC.

127. T. E. Cheatham and P. A. Kollman, Molecular dynamics simulation of nucleic acids. *Annual Review of Physical Chemistry*, (2000) **51**: 435-471.

128. D. Bassolino-Klimas, R. Tejero, S. R. Krystek, W. J. Metzler, G. T. Montelione, and R. E. Bruccoleri, Simulated annealing with restrained molecular dynamics using a flexible restraint potential: Theory and evaluation with simulated NMR constraints. *Protein Science*, (1996) **5**(4): 593-603.

129. W. J. Metzler, D. R. Hare, and A. Pardi, Limited sampling of conformational space by the distance geometry algorithm: implications for structures generated from NMR data. *Biochemistry*, (1989) **28**(17): 7045-7052.

130. A. T. Brunger, P. D. Adams, and L. M. Rice, New applications of simulated annealing in x-ray crystallography and solution NMR. *Structure*, (1997) **5**(3): 325-336.

131. S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi, Optimization by simulated annealing. *Science*, (1983) **220**(4598): 671-680.

132. C. D. Schwieters and G. M. Clore, The VMD-XPLOR visualization package for NMR structure refinement. *Journal of Magnetic Resonance*, (2001) **149**(2): 239-244.

133. E. G. Stein, L. M. Rice, and A. T. Brunger, Torsion-angle molecular dynamics as a new efficient tool for NMR structure calculation. *Journal of Magnetic Resonance*, (1997) **124**(1): 154-164.

134. B. M. Znosko, T. W. Barnes, III, T. R. Krugh, and D. H. Turner, NMR studies of DNA single strands and DNA:RNA hybrids with and without 1-propynylation at C5 of oligopyrimidines. *Journal of the American Chemical Society*, (2003) **125**(20): 6090-6097.

135. Y. Wu, D. Bhattacharyya, C. L. King, I. Baskerville-Abraham, S.-H. Huh, G. Boysen, J. A. Swenberg, B. Temple, S. L. Campbell, and S. G. Chaney, Solution structures of a DNA dodecamer duplex with and without a cisplatin 1,2-d(GG) intrastrand cross-link: comparison with the same DNA duplex containing an oxaliplatin 1,2-d(GG) intrastrand cross-link. *Biochemistry*, (2007) **46**(22): 6477-6487.

136. J. W. Pham, I. Radhakrishnan, and E. J. Sontheimer, Thermodynamic and structural characterization of 2'-nitrogen-modified RNA duplexes. *Nucleic Acids Research*, (2004) **32**(11): 3446-3455.

137. X. Y. Yang, X. G. Han, C. Cross, S. Bare, Y. Sanghvi, and X. L. Gao, NMR structure of an antisense DNA:RNA hybrid duplex containing a 3'-$CH_2N(CH_3)$-O-5' or an MMI backbone linker. *Biochemistry*, (1999) **38**(39): 12586-12596.

138. O. Mauffret, G. Tevanian, and S. Fermandjian, Residual dipolar coupling constants and structure determination of large DNA duplexes. *Journal of Biomolecular NMR*, (2002) **24**(4): 317-328.

139. S. J. Headey, H. Huang, J. K. Claridge, G. A. Soares, K. Dutta, M. Schwalbe, D. Yang, and S. M. Pascal, NMR structure of stem-loop D from human rhinovirus-14. *RNA*, (2007) **13**(3): 351-360.

140. J. Lee, V. Guelev, S. Sorey, D. W. Hoffman, and B. L. Iverson, NMR structural analysis of a modular threading tetraintercalator bound to DNA. *Journal of the American Chemical Society*, (2004) **126**(43): 14036-14042.

141. C. Blanchet, M. Pasi, K. Zakrzewska, and R. Lavery, CURVES[+] web server for analyzing and visualizing the helical, backbone and groove parameters of nucleic acid structures. *Nucleic Acids Research*, (2011) **39**(Web Server): W68-W73.

142. S. Arnott and D. W. L. Hukins, Refinement of the structure of B-DNA and implications for the analysis of x-ray diffraction data from fibers of biopolymers. *Journal of Molecular Biology*, (1973) **81**(2): 93-105.

143. S. Arnott and D. W. Hukins, Optimised parameters for A-DNA and B-DNA. *Biochemical and Biophysical Research Communications*, (1972) **47**(6): 1504-1509.

144. R. R. Sinden, *DNA Structure and Function*. 1st Edition ed. (1994): Academic Press.

145. M. A. El Hassan and C. R. Calladine, Two distinct modes of protein-induced bending in DNA. *Journal of Molecular Biology*, (1998) **282**(2): 331-343.

146. H. R. Drew, R. M. Wing, T. Takano, C. Broka, S. Tanaka, K. Itakura, and R. E. Dickerson, Structure of a B-DNA dodecamer: conformation and dynamics. *Proceedings of the National Academy of Sciences*, (1981) **78**(4): 2179-2183.

147. M. Egli, G. Minasov, M. Teplova, R. Kumar, and J. Wengel, X-ray crystal structure of a locked nucleic acid (LNA) duplex composed of a palindromic 10-mer DNA strand containing one LNA thymine monomer. *Chemical Communications*, (2001)(7): 651-652.

148. http://chem.rutgers.edu/~xiangjun/3DNA/images/bp_step_hel.gif.

149. S. B. Nabuurs, C. A. E. M. Spronk, G. Vriend, and G. W. Vuister, Concepts and tools for NMR restraint analysis and validation. *Concepts in Magenetic Resonance*, (2004) **22**(2): 90-105.

150. A. Arcella, G. Portella, M. L. Ruiz, R. Eritja, M. Vilaseca, V. Gabelica, and M. Orozco, Structure of triplex DNA in the gas phase. *Journal of the American Chemical Society*, (2012) **134**(15): 6596-6606.

151. R. Macaya, E. Wang, P. Schultze, V. Sklenar, and J. Feigon, Proton nuclear magnetic resonance assignments and structural characterization of an intramolecular DNA triplex. *Journal of Molecular Biology*, (1992) **225**(3): 755-773.

152. J. Feigon, K. M. Koshlap, and F. W. Smith, ^{1}H NMR spectroscopy of DNA triplexes and quadruplexes. *Nuclear Magnetic Resonance and Nucleic Acids*, (1995) **261**: 225-255.

153. V. Sklenar and J. Feigon, Formation of a stable triplex from a single DNA strand. *Nature*, (1990) **345**(6278): 836-838.

154. C. Glemarec, R. C. Reynolds, P. A. Crooks, J. A. Maddry, M. S. Akhtar, J. A. Montgomery, J. A. Secrist Iii, and J. Chattopadhyaya, Conformational studies

of thymidine dimers containing sulfonate and sulfonamide linkages by NMR spectroscopy. *Tetrahedron*, (1993) **49**(11): 2287-2298.

155.   S. H. Chou, P. Flynn, and B. Reid, Solid-phase synthesis and high resolution NMR studies of two synthetic double-helical RNA dodecamers - r(CGCGAAUUCGCG) and r(CGCGUAUACGCG). *Biochemistry*, (1989) **28**(6): 2422-2435.

156.   C. Y. Sekharudu, N. Yathindra, and M. Sundaralingam, Molecular-dynamics investigations of DNA triple-helical models - unique features of the Watson-Crick duplex. *Journal of Biomolecular Structure & Dynamics*, (1993) **11**(2): 225-244.

157.   K. M. Koshlap, P. Schultze, H. Brunar, P. B. Dervan, and J. Feigon, Solution structure of an intramolecular DNA triplex containing an N-7-glycosylated guanine which mimics a protonated cytosine. *Biochemistry*, (1997) **36**(9): 2659-2668.

158.   E. Wang, K. M. Koshlap, P. Gillespie, P. B. Dervan, and J. Feigon, Solution structure of a pyrimidine purine pyrimidine triplex containing the sequence-specific intercalating non-natural base D-3. *Journal of Molecular Biology*, (1996) **257**(5): 1052-1069.

159.   D. Leitner, W. Schroder, and K. Weisz, Influence of Sequence-Dependent Cytosine Protonation and Methylation on DNA Triplex Stability *Biochemistry*, (2000) **39**(19): 5886-5892.

160.   A. M. Soto, J. Loo, and L. A. Marky, Energetic Contributions for the Formation of TAT/TAT, TAT/CGC$^+$, and CGC$^+$/CGC$^+$ Base Triplet Stacks. *Journal of the American Chemical Society*, (2002) **124**(48): 14355-14363.

161.   A. Jain, M. R. Rajeswari, and F. Ahmed, Formation and thermodynamic stability of intermolecular (RRY) DNA triplex in GAA/TTC repeats associated with Freidreich's ataxia. *Journal of Biomolecular Structure & Dynamics*, (2002) **19**(4): 691-699.

162.   H.-T. Lee, S. Arciniegas, and L. A. Marky, Unfolding thermodynamics of DNA pyrimidine triplexes with different molecularities. *The Journal of Physical Chemistry B*, (2008) **112**(15): 4833-4840.

163.   J. S. Sun, J. L. Mergny, R. Lavery, T. Montenay-Garestier, and C. Helene, Triple helix structures: sequence dependence, flexibility and mismatch effects. *Journal of Biomolecular Structure & Dynamics*, (1991) **9**(3): 411-424.

164. D. A. Rusling, G. Peng, N. Srinivasan, K. R. Fox, and T. Brown, DNA triplex formation with 5-dimethylaminopropargyl deoxyuridine. *Nucleic Acids Research*, (2009) **37**(4): 1288-1296.

165. T. Michel, F. Debart, J. J. Vasseur, F. Geinguenaud, and E. Taillandier, FTIR and UV spectroscopy studies of triplex formation between alpha-oligonucleotides with non-ionic phoshoramidate linkages and DNA targets. *Journal of Biomolecular Structure & Dynamics*, (2003) **21**(3): 435-445.

166. M. Koizumi, K. Morita, M. Daigo, S. Tsutsumi, K. Abe, S. Obika, and T. Imanishi, Triplex formation with 2'-O,4'-C-ethylene-bridged nucleic acids (ENA) having C3'-endo conformation at physiological pH. *Nucleic Acids Research* (2003) **31**(12): 3267-3273.

167. S. M. A. Rahman, S. Seki, S. Obika, S. Haitani, K. Miyashita, and T. Imanishi, Highly stable pyrimidine-motif triplex formation at physiological pH values by a bridged nucleic acid analog. *Angewandte Chemie, International Edition*, (2007) **46**(23): 4306-4309.

168. A. P. G. Beevers, K. J. Fettes, G. Sabbagh, F. K. Murad, J. R. P. Arnold, R. Costick, and J. Fisher, NMR and UV studies of 3 '-*S*-phosphorothiolate modified DNA in a DNA: RNA hybrid dodecamer duplex; implications for antisense drug design. *Organic & Biomolecular Chemistry*, (2004) **2**(1): 114-119.

169. E. G. Plum and K. J. Breslauer, Thermodynamics of an intramolecular DNA triple helix: A calorimetric and spectroscopic study of the pH and salt dependence of thermally induced structural transitions. *Journal of Molecular Biology*, (1995) **248**(3): 679-695.

170. R. Owczarzy, 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): 5336-5353.

171. J. L. Mergny and L. Lacroix, Analysis of thermal melting curves. *Oligonucleotides*, (2003) **13**(6): 515-537.

172. L. A. Marky and K. J. Breslauer, Calculating thermodynamic data for transitions of any molecularity from equilibrium melting curves. *Biopolymers*, (1987) **26**(9): 1601-1620.

173. K. J. Breslauer and G. K. A. Michael L. Johnson, Extracting thermodynamic data from equilibrium melting curves for oligonucleotide order-disorder transitions. *Methods in Enzymology*, (1995) **259**: 221-242.

174. A. M. Soto and L. A. Marky, Thermodynamic contributions for the incorporation of GTA triplets within canonical TAT/TAT and C$^+$GC/C$^+$GC

base-triplet stacks of DNA triplexes. *Biochemistry*, (2002) **41**(41): 12475-12482.

175. M. H. Caruthers, Chemical synthesis of DNA and DNA analogs. *Accounts of Chemical Research*, (1991) **24**(9): 278-284.

176. Dionex, *Chromeleon Chromatography Management System*. (2006).

177. *ACD Labs*, in *Spec Manager*. (2009), Advanced Chemistry Development, Inc.

178. T. D. Goddard and D. G. Kneller, *Sparky 3*. (2006): University of California, San Fransisco, USA.

179. N. Foloppe and J. A. D. MacKerell, All-atom empirical force field for nucleic acids: I. Parameter optimization based on small molecule and condensed phase macromolecular target data. *Journal of Computational Chemistry*, (2000) **21**(2): 86-104.

180. http://pymol.org/.