Structural and Biophysical Studies of Nucleic Acids

WIMAL PATHMASIRI
Dissertation presented at Uppsala University to be publicly examined in 113a, B7, Husargatan 3, Uppsala, Tuesday, November 6, 2007 at 10:15 for the degree of Doctor of Philosophy. The examination will be conducted in English.

Abstract

This thesis is based on six research publications concerned with (i) study of the molecular structures and dynamics of modified nucleosides; (ii) investigation of the effect of incorporation of modified nucleosides on the structure of DNA; (iii) examination of the effect of the sugar modifications on the pseudo-aromatic properties (pK_a) of the nucleobases; (iv) analysis of the effect of the CH-π interactions on the relative stability of the DNA-RNA hybrid duplexes. The structural stability of the nucleic acids as well as their behavior in molecular recognition is dominated by hydrogen bonding and stacking interactions beside other non-covalent interactions. Naturally occurring nucleosides are found to have some specific functions. Modifications of nucleic acids, followed by studies of the resulting structural, chemical and functional changes, contribute to an understanding of their role in various biochemical processes, such as catalysis or gene silencing. In papers I-III, analysis of the structures of modified thymidine nucleosides with 1',2'-oxetane or azetidine) and 2',4'-(LNA, 2'-amino LNA, ENA, and Aza-ENA) conformationally constrained sugar moieties, and dynamics of the modified nucleosides by NMR, ab initio, and molecular dynamics simulations are discussed. Based on whether the modification leads to 1',2'- or 2',4'-constrained sugar moieties, it is found that they fall into two distinct categories characterized by their respective internal dynamics of the glycosidic and backbone torsions as well as by their characteristic NE-type (P = 37° ± 27°, Φ_m = 25° ± 18°) for 1',2'-constrained nucleosides, and N-type (P = 19° ± 8°, Φ_m = 48° ± 4°) for 2',4'-constrained systems, respectively. Moreover, each group has different conformational hyperspace accessible. The effect of the incorporation of 1',2'-oxetane locked thymidine nucleoside on the structure and dynamics of the Dickerson-Drew dodecamer, d(CGCGAATTCGCG)_o, determined by NMR, is discussed in the paper IV. It shows that the incorporation of oxetane locked T into the dodecamer has made local structural deformations and perturbation in base pairing, where the modification is included. The modulations of physico-chemical properties of the nucleobases in nucleotides by the C2'-modification of the sugar (paper V), 5'-phosphate group, and the effect of constrained pentofuranosyl moiety (sugar, paper III) have been studied. CH-π interactions between the methyl group of thymidine and the neighboring aromatic nucleobase are shown to increase the relative stability of the DNA-RNA hybrid duplexes over the isosequential RNA-DNA duplexes or vice versa (paper VI).

Keywords: nucleic acids, Nuclear Magnetic Resonance, molecular dynamics, sugar modified nucleosides, pKa

Wimal Pathmasiri, Department of Biochemistry and Organic Chemistry, BMC, Box 576, Uppsala University, SE-75123 Uppsala, Sweden

© Wimal Pathmasiri 2007

ISSN 1651-6214
urn:isbn:se:uu:diva-8245 (http://urn.kb.se/resolve?urn=urn:isbn:se:uu:diva-8245)
To my parents and family
THE ORIGINAL PUBLICATIONS

This thesis is based on the following original publications referred by
the roman numerals.


Reprints were made with permission from the publishers.
The author wishes to clarify his contributions to the research presented in the present thesis.

Paper I and II: Acquired all the high-resolution NMR spectra. Performed the detailed NMR characterization of the final compounds, all of the kinetics study in paper I, by high resolution NMR spectroscopy, and spin simulations. Participated in the discussions regarding the manuscript and the part of the manuscript regarding the NMR characterization was written. Documented all supplementary materials related to the NMR characterization by high-resolution NMR spectroscopy.

Paper III: Performed ca. 50% of the pKₐ measurements by NMR-based titrations. Interpreted the data and documented supplementary material together with the co-workers. Involved in the discussions regarding geometry optimization, the molecular dynamics study, and the manuscript preparation.

Paper IV: Contributed ca. 50% to the DNA melting study by NMR and preparation of the related supplementary data and figures for the manuscript. Performed all the NOESY and ROESY NMR experiments in the DNA hydration study, interpreted the data together with the co-workers. Performed the whole T₁ relaxation study and line width analysis by NMR, and interpreted all the data.

Paper V: Discussed the project layout with the supervisor together with the co-workers, performed ca. 40% of the NMR titrations, involved in all the discussions and partly for writing the manuscript, and documented supplementary material.

Paper VI: Performed the data analysis of the hybrid duplexes together with the co-worker. Involved in the discussions regarding the manuscript. Documented the supporting information together with the co-worker.
Contents

1. Nucleic Acids ................................................................................................. 11
   1.1 Composition of DNA and RNA ................................................................. 12
   1.2 Structural properties of nucleic acid components ...................................... 13
       1.2.1 Pseudorotation in pentafuranose sugar ........................................... 13
       1.2.2 Stereoelectronic effects ...................................................................... 15
       1.2.3 Sugar-phosphate backbone torsions ............................................. 15
   1.3 DNA and RNA secondary structure ......................................................... 16
   1.4 Non-covalent interactions stabilizing nucleic acid secondary structure .... 18
       1.4.1 Hydrogen bonding ......................................................................... 19
       1.4.2 Stacking ......................................................................................... 20
       1.4.3 CH-π Interactions .......................................................................... 20
       1.4.4 Hydration of nucleic acids .............................................................. 21
   1.5 Exploration of nucleic acid structure by NMR spectroscopy .......... 22
       1.5.1 Biomolecular NMR ....................................................................... 22
       1.5.2 Structure and dynamics of nucleic acids ....................................... 23

2. Sugar modifications in nucleosides ............................................................... 26
   2.1 Characterization and conformational analysis of modified nucleosides (papers I-III) ................................................................. 26
       2.1.1 Characterization of modified nucleosides ...................................... 27
       2.1.2 Determination of structure and dynamics of 1',2'- and 2',4'- constrained nucleosides .............................................................. 30
   2.2 Effect of sugar modification on the stability of DNA duplex: an NMR study (paper IV) ................................................................. 31
       2.3 Effect of the sugar modification on the physico-chemical properties of the nucleotides (paper III & V) .............................................. 36

3. Effect of CH-π interactions on stability of the hetero-duplexes (Paper VI) ................................................................. 42

4. Conclusion and outlook .................................................................................. 45

5. Acknowledgements ....................................................................................... 47

6. Summary in Swedish ..................................................................................... 50
7. References .................................................................................................................. 53
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1D</td>
<td>One dimensional</td>
</tr>
<tr>
<td>2D</td>
<td>Two dimensional</td>
</tr>
<tr>
<td>3D</td>
<td>Three dimensional</td>
</tr>
<tr>
<td>4D</td>
<td>Four dimensional</td>
</tr>
<tr>
<td>AON</td>
<td>Antisense oligonucleotide</td>
</tr>
<tr>
<td>Aza-ENA</td>
<td>2'-amino ethylene bridged nucleic acids</td>
</tr>
<tr>
<td>B</td>
<td>Base</td>
</tr>
<tr>
<td>CD</td>
<td>Circular dichroism</td>
</tr>
<tr>
<td>CORIMA</td>
<td>Complete relaxation matrix approach</td>
</tr>
<tr>
<td>COSY</td>
<td>Correlation spectroscopy</td>
</tr>
<tr>
<td>DNA</td>
<td>Deoxyribonucleic acid</td>
</tr>
<tr>
<td>DD</td>
<td>DNA-DNA</td>
</tr>
<tr>
<td>DQF-COSY</td>
<td>Double quantum filtered correlation spectroscopy</td>
</tr>
<tr>
<td>DR</td>
<td>DNA-RNA</td>
</tr>
<tr>
<td>E-type</td>
<td>East-type</td>
</tr>
<tr>
<td>$E_a$</td>
<td>Activation energy</td>
</tr>
<tr>
<td>ENA</td>
<td>Ethylene bridged nucleic acids</td>
</tr>
<tr>
<td>H-Bond</td>
<td>Hydrogen bond</td>
</tr>
<tr>
<td>HMBC</td>
<td>Heteronuclear multiple bond coherence</td>
</tr>
<tr>
<td>HMQC</td>
<td>Heteronuclear multiple quantum coherence</td>
</tr>
<tr>
<td>HOMO</td>
<td>Highest occupied molecular orbital</td>
</tr>
<tr>
<td>ISPA</td>
<td>Isolated spin pair approximation</td>
</tr>
<tr>
<td>LC</td>
<td>Liquid chromatography</td>
</tr>
<tr>
<td>LNA</td>
<td>Locked nucleic acid</td>
</tr>
<tr>
<td>ON</td>
<td>Oligo Nucleotide</td>
</tr>
<tr>
<td>Oxetane T</td>
<td>Oxetane constrained thymidine</td>
</tr>
<tr>
<td>LUMO</td>
<td>Lowest unoccupied molecular orbital</td>
</tr>
<tr>
<td>MD</td>
<td>Molecular dynamics</td>
</tr>
<tr>
<td>mRNA</td>
<td>Messenger RNA</td>
</tr>
<tr>
<td>N-type</td>
<td>North-type</td>
</tr>
<tr>
<td>NE-type</td>
<td>North East-type</td>
</tr>
<tr>
<td>NMR</td>
<td>Nuclear magnetic resonance</td>
</tr>
</tbody>
</table>
NOE  Nuclear Overhauser effect
NOESY  Nuclear Overhauser effect spectroscopy
OMe  Methoxy
ON  Oligonucleotide
P  Phase angle
P.E. COSY  Primitive exclusively correlation spectroscopy
Pu  Purines
Py  Pyrimidines
RD  RNA-DNA
RMSd  Root mean square deviation
RNA  Ribonucleic acid
RNAi  RNA interference
ROESY  Rotating frame Overhauser effect spectroscopy
rRNA  Ribosomal RNA
siRNA  Small interfering RNA
SA  Simulated annealing
S-type  South-type
T  Thymine
tRNA  Transfer RNA
TOCSY  Total correlation spectroscopy
TROSY  Transverse relaxation optimized spectroscopy
W-type  West-type
WC  Watson Crick
ΔG‡  Free energy change of activation
ΔG  Free energy change of a process
ΔH  Enthalpy change of a process
ΔS  Entropy change of a process
1. Nucleic Acids

The nucleic acids are large biological polymers. They naturally exist as one of two similar chemical forms: Deoxyribonucleic acid (DNA) and Ribonucleic acid (RNA). Monomer units of nucleic acids are nucleotides, each of which is the phosphate esters of nucleosides. Generally, the monomer units of DNA or RNA consist of four different nucleosides. The length of nucleic acids can vary from 80 in tRNA to some $10^8$ nucleotides in chromosomes. Nucleic acids are the molecular repositories for genetic information. The structure of every protein, and ultimately in every cellular constituent is the result of the coded information programmed into the nucleic acids. In addition, cells contain many types nucleotides, which have a wide variety of cellular processes, including metabolic regulation.

DNA serves as the storage and carrier units of genetic information. Several classes of cellular RNA are found: messenger RNA (mRNA) serves as the intermediary between the storage and translation; transfer RNA (tRNA) serves as transporters of amino acids, the building blocks of proteins, in the translation process; ribosomal RNA (rRNA) is a structural component of the ribosome, the translation machinery. Apart from this, RNA is involved in different functions including self-splicing and catalysis (Ribozyme). DNA is also involved in other functions such as DNAzyme in catalysis. It has been reported that there are about 93 post-transcriptionally modified nucleosides found in RNA. These modifications may be found in the nucleobase, sugar, or in the phosphate unit and have long been recognized as characteristics of RNA and have specific functions such as codon recognition. The ribonucleotides have a 2'-OH group, which is involved in important biological functions such as RNA splicing and hydrogen bonding.

Chemically modified nucleic acids function as model systems for native DNA and RNA; as chemical probes in diagnostics or analysis of protein-nucleic acid interactions and in high throughput genomics and drug target validation; as potential antigen-, antisense, or RNAi based drugs; and as tools for structure determination; as tools for understanding mechanisms of RNA and DNA catalysis (Ribozyme and DNAzyme) and cell signaling processes. These modifications can be in the phosphate linkages (e.g., phosphorothioate, phosphorodithioate, methylphosphonate), sugar moieties (C1',C2'-constrained, C2',C4'-constrained, C2', or O4'), or in the nucleobases (purine, pyrimidine, abasic modification). C2'-modification of
the sugar moiety is widely used to make oligonucleotides thermodynamically and nucleolytically stable as well as to recruit RNase H enzyme in the antisense approach for the down-regulation of gene expression. They are also found to be extremely useful in the design of unique binding properties of aptamers to specific ligands by in vitro evolution.

Incorporation of these modified molecules into the nucleic strands have also found to be promising in terms of target binding, accessibility, and nuclease resistance. The incorporation of the modification to the native nucleic acids undoubtedly affects their structure. The study of the effect of incorporation of modified nucleosides on their structure shed the light for better understanding their functions. In addition, the sugar modifications have direct effects on the properties of the nucleobases such as donor acceptor properties in base pairing.

The understanding of the structural and physico-chemical properties of these modified nucleosides and nucleic acids provides opportunities for optimizing better target molecules. Hence, this thesis is a study of: (i) structural and conformational analysis of modified nucleosides; (ii) solution structure and dynamics of DNA by NMR spectroscopy; (iii) effects of C2' substituent of the sugar on the pseudoaromatic character of nucleobase; (iv) effect of CH-π interactions on the stability of the heteroduplexes.

1.1 Composition of DNA and RNA

Nucleotides have three components: a heterocyclic nitrogenous base, a pentafuranose sugar, and a phosphate group (Figure 1). In the nucleotides, the bases are either purines or pyrimidines. A comparison of DNA and RNA characteristics are shown in Table 1. A number of modified nucleobases are also found in nature, for example, in tRNA. The C1’ of the sugar is linked to the N9 of purines or N1 of pyrimidines by glycosidic bonds. The base is attached to the sugar in the β configuration (i.e. base and the C5' are on the same side).

<table>
<thead>
<tr>
<th>Parameter</th>
<th>DNA</th>
<th>RNA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pyrimidines</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Attachment of sugar (C1’) to base</td>
<td>N9 of purine, N1 of pyrimidine</td>
<td>N9 of purine, N1 of pyrimidine</td>
</tr>
<tr>
<td>Sugar</td>
<td>β-D-2’-deoxyribose</td>
<td>β-D-ribose</td>
</tr>
<tr>
<td>Phosphate attachment sites</td>
<td>3’, 5’</td>
<td>2’, 3’, 5’</td>
</tr>
<tr>
<td>Conformational flexibility</td>
<td>higher</td>
<td>lower</td>
</tr>
<tr>
<td>Sensitivity to hydrolysis</td>
<td>lower</td>
<td>higher</td>
</tr>
</tbody>
</table>

Table 1. Comparison of DNA and RNA characteristics.
Figure 1. General structure of nucleic acids, exemplified by the nucleotide sequence AGCU(T). Each nucleotide consists of a base, a pentose sugar, and a phosphate group. DNA: R = H, the bases are A, G, C, and T. RNA: R = OH, the bases are A, G, C, and U. One nucleoside is marked by brackets.

1.2 Structural properties of nucleic acid components

Some of the most important parameters used to characterize the structural properties of nucleic acids are summarized as follows.

1.2.1. Pseudorotation in pentafuranose sugar

It has long been recognized\(^{36}\) that the furanose ring of the nucleosides or nucleotides is not planar but puckered. A planar five-membered sugar ring is sterically and thus energetically unfavourable\(^{37, 38}\). The intra-molecular ring strain is therefore released by moving one atom out of the ring plane\(^{1-3, 38}\). Consequently, the conformational energy is lowered and a stable sugar con-
formation is formed. The many possible puckered states are best described in the framework of a pseudorotation cycle\(^1\textsuperscript{3},\textsuperscript{39} \textsuperscript{44}\) (Figure 2).

In contrast to cyclopentane, where the pseudorotation is free, the pseudorotation in nucleic acid constituents is hindered\(^3\) because of their substituents. The preferred puckered states involve either C3' or C2'. The conformation with C2' carbon out of the plane (C4'-O4'-C1') and on the same side as C5' is called C2'-endo (\(^3E\)) whereas the conformation is called C3'-endo (\(^3E\)) when the C3' is out of plane instead. The conformation can be either pure C2'-endo/ C3'-endo forms or mixed C2'-endo-C3'-exo/ C2'-exo-C3'-endo forms. Other conformations such as pure C2'-exo, C3'-exo, C4', O4', C1'-endo or –exo are also possible.

Two parameters\(^3,\textsuperscript{40},\textsuperscript{42},\textsuperscript{45}\), the phase angle (P) and the puckering amplitude (\(\Phi_m\)), are used to describe the endocyclic torsion angles, \(\nu_0 - \nu_4\), in the pseudorotation cycle, according to the formula: \(\nu_j = \Phi_m \cos [P + 144 (j-2)]\), \(j=0,1,2,3,4\). The forms C3'-endo-C2'-exo (\(P \approx 0–18^\circ\), N-type) and the C3'-exo-C2'-endo (\(P \approx 162-180^\circ\), S-type) conformations are the most populated sugar conformations observed\(^3\) in RNA and DNA, respectively. Usually, \(\Phi_m\) varies between 34° and 42°. The energy barrier of N-type and S-type equilibrium is 1.2–5\(^3\textsuperscript{8}\) kcalmol\(^{-1}\). There are a few E-type conformations\(^46\) observed whereas no W-type is possible\(^47\) due to the steric effects.

![Figure 2](image.png)

*Figure 2.* Pseudorotation cycle of the nucleoside sugar moiety. Shaded areas show the most abundant sugar conformations found in nucleic acids.

Due to the relatively low energy barrier, the sugar conformation is driven into rapid transition of the two low energy conformers and often treated as a two state equilibrium of the N/S conformations. The average sugar conformation is determined by the balance of stereoelectronic effects\(^3\textsuperscript{8},\textsuperscript{48}\), i. e., anomeric and gauche effects.
1.2.2. Stereoelectronic effects

A nucleobase promoted anomeric effect\(^ {38, 49-53}\) drives the sugar towards the \(N\)-type conformations, where the stabilizing \(\sigma_{\text{C1'-N1/N9}}\) orbital mixing is feasible\(^ {38}\). The anomeric effect in DNA is balanced by steric and gauche effects,\(^ {38, 50, 54, 55}\) which drive the sugar towards the \(S\)-type conformation. There is less steric effect between the sugar and the base when the base is in pseudo-equatorial orientation, which is the case in \(S\)-type conformation. Gauche effect stabilizes the \(S\)-type conformation through the overlap of \(\sigma_{\text{C1'-H3'}}\) orbitals. In addition, in RNA, the 2'-OH creates other gauche interactions as well, i.e. O2' \(\rightarrow\) O4' and O2' \(\rightarrow\) N9/ N1 overlaps\(^ {52}\). The overall effect of these interactions drives the sugar conformation towards the \(N\)-type. Moreover, the \(N\)-\(S\) equilibrium can be steered towards either direction by: (i) the nature of the nucleobase\(^ {38, 48}\) (electron density); (ii) electronegative substitution in the sugar\(^ {38, 48}\); (iii) environmental factors such as salt, relative humidity, solvent, or nucleobase sequence.

1.2.3 Sugar-phosphate backbone torsions

A set of rotations characterizing the nucleotide sugar-phosphate backbone involves six independent angles\(^ {1-3, 44}\) starting from 5'-phosphate group towards the 3'-end: \(\alpha, \beta, \gamma, \delta, \varepsilon,\) and \(\zeta\) (Figure 3). In addition, the base is attached to the C1' of sugar by a glycosidic bond by N1 of a pyrimidine or N9 of a purine. The torsion angle about this bond is specified by the angle \(\chi\).\(^ {2, 3}\) The two ranges of angles found with \(\chi\) are designated as \(\text{syn} (-90^\circ \leq \chi \leq 90^\circ))\) and \(\text{anti} (90^\circ \leq \chi \leq 270^\circ))\).

![Figure 3](image_url)

*Figure 3.* Seven torsion angles are used to define the sugar-phosphate backbone. Endocyclic sugar torsion angles \(\nu_0-\nu_4\) are also shown here.

The rotations around the phosphate-sugar backbone is quite variable, but conformations with bases stacked in a helix tend to have \(\alpha\) (g-), \(\beta\) (t), \(\gamma\) (g+),
ε (t), and ζ (g-) because of the local rigidity and steric interactions among the groups³. The angle δ depends on the extent of sugar puckering. Based on the sugar conformation, C3'-endo or C2'-endo, the distances between P5' and P3' in the nucleic acid chain is different: 5.9 Å and 7.0 Å, respectively.

1.3 DNA and RNA secondary structure

In addition to the first proposed secondary structure of nucleic acids, i.e., Watson and Crick’s right handed double helix of DNA⁵⁶, many additional structures have been discovered¹-³. In the double helix, the bases from the two strands are paired by inter-molecular hydrogen bonds between complementary bases in the interior, the sugar phosphate backbone extends along the outside, and the strands run in anti-parallel directions. DNA is found in single, double², or multiple stranded (triplexes⁵⁷ and quadruplexes⁵⁸) as well as the super-coiled⁴ and hairpin⁵⁹ structures. In addition to the single and double stranded RNA, complex structures (e.g. pseudoknots⁶⁰-⁶²) for tRNA, small nuclear RNA, and catalytic RNA (ribozyme) have also been reported.

Several distinct classes of helical structures of nucleic acids have been identified⁶³-⁶⁵ depending upon their environment such as relative humidity, types of salt, ionic strength, and the solvent. DNA is now known to exist in three different helical forms³,⁶⁶: A-, B-, and Z-forms (Figure 4). The B-form is the dominant form, and B- to A- transition of DNA has also been observed. RNA exists predominantly in A-form. It has also been shown earlier⁶⁷ that single stranded DNA and RNA have a pre-organized helical structure.

<table>
<thead>
<tr>
<th>A-DNA</th>
<th>B-DNA</th>
<th>Z-DNA</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1.png" alt="A-DNA" /></td>
<td><img src="image2.png" alt="B-DNA" /></td>
<td><img src="image3.png" alt="Z-DNA" /></td>
</tr>
</tbody>
</table>

*Figure 4. The different representations of double helical DNA: A, B, and Z-forms, each containing 20 base pairs. Ball and stick side views and cylindrical end views and cylindrical representations are shown here. Reproduced from reference 3, Copyright ©2000 University Science Books, Sausalito, CA.*
A- and B-forms are right-handed helices. Both DNA and RNA can adopt a left-handed Z-form at high salt concentration, and occur in alternating purine (Pu)–pyrimidine (Py) sequences, mainly GC base steps. It is well known that the sugar conformation in the A-form is C3′-endo whereas that of the B-form is C2′-endo. It has been reported for Z-forms that alternating C2′-endo and C3′-endo sugar conformations are found in Pu and Py residues, respectively. The glycosidic torsion anti conformation is energetically more stable than the syn conformation. The syn conformation has been found in G residues in Z-form of DNA and RNA.

A set of helical parameters (translational and rotational, Figure 5 and Table 2) has been defined in order to specify the structural features of different helical forms of DNA and RNA. Among them, x-displacement is shown to best differentiate the three different forms, A-, B- or Z-forms.

Figure 5. Definitions of translations and rotations involving bases and base pairs. The shaded edge represents the minor groove and the black corners are the points of attachment of the glycosidic bonds to the C1′ sugar atoms. Reproduced from reference 3, Copyright ©2000 University Science Books, Sausalito, CA.
Table 2. Average Structural Parameters for Different Helical Forms. Reproduced from reference 3, Copyright ©2000 University Science Books, Sausalito, CA.

<table>
<thead>
<tr>
<th></th>
<th>A-DNA</th>
<th>B-DNA</th>
<th>Z-DNA</th>
<th>Z (WC) DNA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Helical Handedness</td>
<td>Right</td>
<td>Right</td>
<td>Left</td>
<td>Left</td>
</tr>
<tr>
<td>bp/ repeating unit</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>bp/ turn</td>
<td>11</td>
<td>10</td>
<td>12</td>
<td>12</td>
</tr>
<tr>
<td>Helix twist, (°)</td>
<td>32.7</td>
<td>36.0</td>
<td>-10&lt;sup&gt;a&lt;/sup&gt;, -50&lt;sup&gt;b&lt;/sup&gt;</td>
<td>-68&lt;sup&gt;a&lt;/sup&gt;, +8&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>Rise/ bp, (Å)</td>
<td>2.9</td>
<td>3.4</td>
<td>-3.9&lt;sup&gt;a&lt;/sup&gt;, -3.5&lt;sup&gt;b&lt;/sup&gt;</td>
<td>-3.9&lt;sup&gt;a&lt;/sup&gt;, -3.5&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>Helix pitch, (Å)</td>
<td>32</td>
<td>34</td>
<td>45</td>
<td>45</td>
</tr>
<tr>
<td>Base pair inclination, (°)</td>
<td>12</td>
<td>2.4</td>
<td>-6.2</td>
<td>-5.8</td>
</tr>
<tr>
<td>P distance from Helix axis, (Å)</td>
<td>9.5</td>
<td>9.4</td>
<td>6.2&lt;sup&gt;a&lt;/sup&gt;, 7.7&lt;sup&gt;b&lt;/sup&gt;</td>
<td>5.6&lt;sup&gt;a&lt;/sup&gt;, 9.1&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>X displacement from bp to helix axis, (Å)</td>
<td>-4.1</td>
<td>0.8</td>
<td>3.0</td>
<td>-1.6</td>
</tr>
<tr>
<td>Glycosidic bond Orientation</td>
<td>anti</td>
<td>anti</td>
<td>anti&lt;sup&gt;c&lt;/sup&gt;, syn&lt;sup&gt;d&lt;/sup&gt;</td>
<td>anti&lt;sup&gt;c&lt;/sup&gt;, syn&lt;sup&gt;d&lt;/sup&gt;</td>
</tr>
<tr>
<td>Sugar Conformation</td>
<td>C3’-endo</td>
<td>C2’-endo&lt;sup&gt;c&lt;/sup&gt;</td>
<td>C2’-endo&lt;sup&gt;c&lt;/sup&gt;</td>
<td>C3’-endo&lt;sup&gt;d&lt;/sup&gt;</td>
</tr>
<tr>
<td>Major groove depth, Width, (Å)</td>
<td>13.5</td>
<td>8.5</td>
<td>Convex</td>
<td>Flat</td>
</tr>
<tr>
<td>Minor groove depth, Width, (Å)</td>
<td>2.7</td>
<td>11.7</td>
<td>9</td>
<td>Deep</td>
</tr>
</tbody>
</table>

<sup>a</sup>CpG step, <sup>b</sup>GpC step, <sup>c</sup>cytocine, <sup>d</sup>guanine, <sup>c</sup>There is a range of conformations

1.4 Non-covalent interactions stabilizing nucleic acid secondary structure

The weak non-covalent interactions<sup>70-73</sup> have a constitutive role in biological systems such as nucleic acids and proteins as well as non-biological systems. Covalent bonds are usually shorter than 2 Å, while noncovalent interactions function within range of several Ångstroms. There are several non-covalent interactions: ion-pairing, hydrogen bonding<sup>74</sup>, CH-π<sup>75-79</sup>, cation-π<sup>72, 73</sup>, π-π interactions<sup>80, 81</sup>. These comprise interactions between permanent dipoles, between a permanent dipole and an induced dipole, and between a time-variable dipole and an induced dipole. The noncovalent complexes are stabilized by: electrostatic (or Coulombic), induction, charge transfer, and dispersion energy contributions. Induction refers to the general ability of charged molecules to polarize neighboring species, and dispersion (London) interaction results from the interactions between fluctuating dipoles. In charge-transfer interactions the electron flow from the donor to the acceptor is indicated.
The three constituent moieties in the nucleotides contribute to the stability of nucleic acids. The phosphate groups in the backbone make the chain a polyelectrolyte. It can interact with different ions as well as make hydrogen bonds with water. The sugar moiety gives the flexibility and the nucleobase is involved in the self-assembly and the molecular recognition processes. It is, however, not clear whether the sugar drives the phosphate geometry or the phosphate drives the sugar conformations in the chain.

The major non-covalent interactions contributing to the self-organization processes in the DNA and RNA are: (i) Intermolecular hydrogen bonding between complementary base pairs; (ii) intra- and inter-residue hydrogen bonding of 2'-OH (in RNA); (iii) stacking; (iv) CH-π interactions; (v) stereoelectronic effects; (vi) hydration in the minor and major grooves.

1.4.1 Hydrogen bonding

Hydrogen bonding is one of the most important non-covalent interactions in biomolecules that regulate various functions in living cells such as stability and molecular recognition. Hydrogen bonds can be categorized as strong (15-40 kcal mol$^{-1}$), moderate (4-15 kcal mol$^{-1}$), or weak (< 4 kcal mol$^{-1}$). For a hydrogen bond to be formed, there should be a donor (AH), which is basically a Lewis acid, and an acceptor atom (B), a Lewis base. Three scalar quantities are necessary to define the geometry of the hydrogen bond: A–H covalent bond length; H---B hydrogen bond length; A---B hydrogen bond distance. This angle is ~ 180°. Several possible types are available, i.e. two centered, three centered, and in some occasions even four centered.

The strength of hydrogen bonding is regulated by the acid-base behavior of the functional groups involved. The difference in the p$K_a$ values for the donor and the acceptor is a measure of the strength of hydrogen bonding. The hydrogen bond strength is at maximum when the p$K_a$ of the donor and acceptor is matched, i.e. Δp$K_a$=0. It is also suggested that, when a hydrogen bond is formed between groups of similar p$K_a$s, the orbital overlap takes place between their matched energy states and results in partial covalent character.

Hydrogen bonding plays a major role in the self-organization and molecular recognition of nucleic acids. Watson-Crick (WC) type of base pairing is predominant whilst other types such as reverse WC, Hoogsteen-type, reverse-Hoogsteen, wobble, and reverse wobble base pairs can also be seen. It has been shown that the base pairing in RNA is stronger compared to that of DNA on the basis of Δp$K_a$ values for rA:rU and rG:rC over dA:dT and dG:dC, which have also been evidenced in literature by deuterium isotope effects using TROSY-HSQC experiments, and by imino proton exchange rates of GC flanking AU(T).
1.4.2 Stacking
Inter- and intra-molecular \( \pi-\pi \) stacking interactions\(^{81, 100-102} \) are a major type of non-covalent interactions involved in both biological and non-biological systems on their stability, molecular recognition, and other functions. These interactions can be categorized according to their geometry: face-face, edge-face, and offset stacking (Figure 6). Stacking interactions provide a major contributive force for the stability of nucleic acid helices\(^3, 80, 84, 89, 92, 103-106 \).

Edge-to-face Offset Face-to-face

*Figure 6. Basic types of \( \pi-\pi \) stacking patterns.*

1.4.3 CH-\( \pi \) Interactions
CH-\( \pi \) interactions\(^{75} \) play an important role in many fields such as in crystal formation, organic reactions, conformational equilibria, and molecular recognition. They occur between an aliphatic CH (soft acid) and an aromatic \( \pi \) (soft base) system\(^{78} \). It is considered as the weakest extreme of hydrogen bonds\(^{75, 78, 107} \). Theoretical calculations have confirmed its borderline character\(^{75} \) with Van der Walls complexes. A number of studies\(^{76-78} \) have shown that CH/\( \pi \) interactions are involved in the stability of biomolecules. Several theoretical studies have also shown that the stabilization of CH-\( \pi \) systems originates in the dispersion energy\(^{75} \). To be effectively participating in CH-\( \pi \) interactions, the H on CH should be positioned above (and perpendicular to) the plane of the aromatic system (Figure 7).
Figure 7. CH-π interactions found in a duplex. A part of the crystal structure of CGCGAATTC⁺CCG (Yoji Umezawa, et. al., Thymine-methyl/π implicated the sequence dependant deformability of DNA, *Nucleic Acids Research*, **2001**, *30*, 2183-2192, by permission of Oxford University Press)

1.4.4 Hydration of nucleic acids

The interaction of water molecules around nucleic acids can be very important in stabilizing helix structure².³ Many water molecules per base pair of DNA can be seen in X-ray structures⁸⁵, ¹⁰⁸-¹¹¹. In the best-resolved structures, up to 14 water molecules per base pair have been observed¹. Highly ordered water molecules can be seen in both minor and major grooves⁸⁵ (Figure 8). The C=O, N, and NH functions of nucleobases as well as phosphate groups interact with water molecules¹.

Figure 8. The minor groove of the d(CGCGAATTCGCG)₂, depicting the arrangement of the water molecules. The fused four water hexagons are shown. Reproduced from Egli, M. et. al. *J. Am. Chem. Soc.* **1999**, *121*, 15, Copyright ©1999 American Chemical Society.
1.5 Exploration of nucleic acid structure by NMR spectroscopy

Nuclear Magnetic Resonance (NMR) spectroscopy represents the most versatile and informative spectroscopic technique used in the chemical research laboratory\textsuperscript{112}. It has a vast number of potential applications ranging from structure elucidation of small molecules to the structure determination of macromolecules. With the recent advancement in both the hardware and in multi-pulse techniques including multidimensional spectroscopy, there has been an increased interest for NMR spectroscopy in the fields of metabolomics\textsuperscript{113-117}, metabonomics\textsuperscript{116, 118-121}, and structural genomics\textsuperscript{122}, more recently also in combination with so called hyphenated techniques\textsuperscript{123-126} such as LC-NMR or LC-NMR-MS\textsuperscript{127}. Important improvements of NMR hardware include, besides the development of high-field magnets, the introduction of micro-coils, capillary-NMR, and cryogenic probes, which have dramatically increased the sensitivity of this technique\textsuperscript{128, 129}.

1.5.1 Biomolecular NMR

Many studies of biomolecular function rely on the knowledge of structure. Since biomolecules typically carry out their function in solution, their dynamics are also important. The two methods available that are capable of three dimensional structure determinations at atomic resolution are X-Ray diffraction and NMR spectroscopy\textsuperscript{3}. In X-Ray diffraction, the molecules exist in a highly ordered crystal lattice, basically trapped in a single conformation, and the structure obtained may or may not represent the structure in solution. Compounds in solution can interconvert between different possible structures. In contrast to X-Ray diffraction, NMR spectroscopy is a well-established method for characterizing three-dimensional molecular structure, dynamics, and interactions of biomolecules in solution\textsuperscript{130-137}. Solution NMR techniques make it possible to study the structural and conformational features of biomolecules including proteins and nucleic acids over a wide range of conditions such as different ionic strengths, pH, solvent compositions, and temperatures\textsuperscript{5}.

Structure determination of biomolecules by NMR usually involves two steps\textsuperscript{138}. The first step is the sequential assignment of the chemical shifts whereas the second step involves the deduction of structural information from nuclear Overhauser effects, scalar couplings, and residual dipolar couplings\textsuperscript{139}. One particular challenge with large molecules is resonance overlap. Isotope labeling methods\textsuperscript{140} and the use of multidimensional NMR experiments\textsuperscript{141, 142} (2D, 3D or 4D, such as HSQC-TOCSY, HSQC-NOESY, TROSY-HSQC, TROSY-NOESY\textsuperscript{143} and $^1$H$^1$C$^{31}$P experiments\textsuperscript{144-147}) have overcome this problem.
A number of nuclei occur in nucleic acids, which can be detected in NMR: $^1\text{H}$, $^{31}\text{P}$, $^{13}\text{C}$, and $^{15}\text{N}$ (all with spin $\frac{1}{2}$). Because of their natural abundance, $^1\text{H}$ and $^{31}\text{P}$ can be detected easily with good sensitivity whereas isotopic labeling methods are needed for detecting $^{13}\text{C}$ and $^{15}\text{N}$. With the recent development of NMR spectroscopy, a number of multi-pulse 1D methods such as 1D NOE and 1D TOCSY are also available to obtain information efficiently within short times. The coupling constants are used to derive dihedral angles involving $^1\text{H}$, $^{13}\text{C}$, and $^{31}\text{P}$ nuclei in the sugar-phosphate backbone in nucleic acids.

1.5.2 Structure and dynamics of nucleic acids

1.5.2.1 Structural features in DNA and RNA

As mentioned earlier (Ch. 1.2.3), seven dihedral angles ($\alpha$, $\beta$, $\gamma$, $\delta$, $\varepsilon$, $\zeta$, and $\chi$, Figure 3, p.9) and the extent of sugar puckering ($P$, Figure 2, p.8) are needed to describe the geometry of nucleic acids. DNA and RNA can exist in single, double, or multiple stranded forms. Furthermore, right- or left-handed helical forms such as A-, B-, or Z-forms of helices are possible. Hence the goal of using NMR in structure determination is to determine these parameters in solution.

1.5.2.2 Structure determination by NMR

A series of 1D, 2D, and 3D NMR spectra are usually employed in the assignment of resonances and the evaluation of constraints for structure calculation.

Assignment strategies

The sequential assignment of protons in helical structures is primarily based on using inter- and intra-residue NOESY cross-peaks between non-exchangeable as well as exchangeable protons. Correlation spectroscopy in terms of TOCSY and COSY spectra is used for the intra-residue resonance assignment of sugar protons.

Nuclear Overhauser Effect

The nuclear Overhauser effect (NOE) is a result of redistribution of spin populations. Such redistributions occur as a result of longitudinal spin relaxation (via dipolar interactions). To induce the spin transitions, a molecule should tumble or rotate with proper frequency similar to the difference between energy levels. The rate at which a molecule tumbles or rotates in solution is defined by the correlation time, $\tau_c$. This means that rapidly tumbling molecules have short correlation times whilst slowly tumbling molecules have long correlation times. A very rough estimate of the correlation...
time of a molecule with particular molecular weight can be obtained from the relationship\(^{112}\): \(\tau_c \approx M \times 10^{-12}\) s.

Intensities for NOESY cross peaks obtained for different mixing times are quantitatively (using available software) or qualitatively evaluated. The characteristic inter- and intra-residue intensity patterns are used to evaluate to determine the right or left-handedness and A-, B-, or Z forms of DNA/RNA as well as the syn or anti configurations alongside the glycosidic bond\(^ {153}\).

**Distance constraints**

NOESY data are routinely used to determine the three dimensional structure of biomolecules. In the first approximation, the isolated spin pair approximation (ISPA)\(^ {150}\), the NOE build up rate is related to the inverse sixth power of the inter-proton distance: \(r_{ij} = r_0(R_0/R_{ij})^{1/6}\), where \(r_{ij}\) and \(R_{ij}\) are the distance and the NOE build up rate between the two protons i and j, respectively. \(R_0\) and \(r_0\) are the distance and the buildup rate for reference protons (for example, \(r_0, H6-H5\) of pyrimidine = 2.54 Å. Buildup rates can be calculated by acquiring NOESY spectra with different mixing times. A number of factors, other than ISPA, contribute to the NOE intensities: the time the magnetization transfer is allowed (mixing time), molecular motion (correlation time), spin diffusion, and base line distortions\(^ {150}\).

As the mixing time increases, the cross-peak intensity in NOESY spectra ceases to develop linearly as a result of spin diffusion or relaxation\(^ {150}\). The linear relationship of inter-proton distance with cross-peak intensities is straightforward only in the linear buildup regime (short mixing time) and for a molecule tumbling as a rigid body. In order to overcome the limitations mentioned above, two methods are generally used\(^ {150}\): Complete Relaxation Matrix Analysis and semi-quantitative distance estimation. The Complete Relaxation Matrix Analysis method relies on the explicit solution of the complete matrix of interactions describing all dipolar interactions. This method does not account satisfactorily for the molecular motion. Therefore, a second method, i.e., the semi-quantitative distance estimation, is widely used. In this method, the distances based on the NOE intensities are classified in loose categories: strong (1.8-3.0 Å), medium (2.5-4.0 Å), and weak (4.0-5.0 Å).

**Dihedral constraints**

\(N\)-type and \(S\)-type sugar populations are found by NMR spectroscopy to be in equilibrium in solution. The vicinal coupling constants between \(H1'\) and \(H2'\) (\(\delta J_{H1'-H2'}\)) can be used to derive possible average sugar conformations by using correlation spectroscopy\(^ {67}\) (P.E. COSY or DQF-COSY) experiments. The \(N\)-type or \(S\)-type sugar populations can be evaluated on the basis of the coupling constants (\(\Sigma J_{H1'}\) for DNA and \(J_{H1'-H2'}\) for RNA). Depending on the resolution and the complexity of the spectrum not all the assignments and...
coupling information are accessible, and therefore it may be required to make use of other methods resolving the resonances, such as 3D NMR spectroscopy\textsuperscript{3, 148, 149} and deuterium labeling\textsuperscript{140} techniques.

Qualitative NMR data can be used to draw some conclusions on the conformational hyperspace of the DNA/ RNA backbone\textsuperscript{154}. For example, P(n)-H4′(n) is only detectable when the four bonds in the H4′-C4′-C5′-O5′-P′ backbone are located in the same plane forming a \textit{W}-shaped configuration. This is only possible when β and γ torsions are in trans and gauche (g+) conformations, respectively, which is the most common conformation for both A- and B-DNA forms. Moreover the vicinal coupling $^3J_{H4H5}$ is very sensitive to the torsion angle γ. When γ is g+ conformation, there will be no strong coupling between H4′ and any of the H5′s (1-2.5 Hz) while both trans and g- will result in a strong coupling the H4′ and any one of the H5′s (10 Hz). If ε is in g- conformation, it produces a detectable coupling between $^4J_{H1′2P}$. The torsion angles, α, β, γ, δ, ε, and ζ, can also be quantitatively estimated by obtaining coupling constant data based on various 3D NMR spectroscopic methods depending on the availability of isotope labeling techniques.

\textit{Watson-Crick Constraints}

In the simulated annealing step, the system is heated and then cooled down to lower temperature. In order to maintain the base pairs intact it is required to include base pair parameters.

\textit{Structure refinement}

After building an initial structure for the sequence, the structure can be refined by the incorporation of experimentally derived distances, dihedral angles, and base pair constraints in the simulated annealing and MD simulation steps to obtain the NMR-refined structures by using a suitable refinement program\textsuperscript{155-157}. Structures are accepted on the results obtained by the evaluation of NMR violations, R- and Q-factor analysis\textsuperscript{158}. 

25
2. Sugar modifications in nucleosides

As mentioned earlier in Chapter 1, modified nucleic acids have a variety of uses. These modifications are found in the nucleobases, sugar, or in the phosphate groups. Sugar modification is a widely used technique in gene silencing, as well as in antisense technology.

In chapter 2.1, structure determination of two modified nucleosides, 2'-N,4'-C-ethylene-bridged thymidine (Aza-ENA T) and 2',4'-carbocyclic ribo-thymidines, the dynamic behavior of 1’, 2’- and 2’, 4’- constrained thymidine nucleosides is discussed. The effect of 1’, 2’-constrained oxetane T nucleoside on the structure and dynamics of the Dickerson-Drew dodecamer is discussed in chapter 2.2. It has been found that the sugar modification as well as the size of the pentafuranose moiety steer the physicochemical properties of the nucleobase, reflected by its pK_a and this is discussed in chapter 2.3.

2.1 Characterization and conformational analysis of modified nucleosides (papers I-III)

In search of conformationally constrained sugar modified nucleosides, the syntheses of 2'-N,4'-C-ethylene-bridged thymidine (Aza-ENA-T), and of five- and six-membered 2’,4’-carbocyclic ribo-thymidines were designed. Prior to further incorporation of these modified nucleosides into oligonucleotides, a complete characterization to confirm the molecular structure and stereochemistry and conformational analysis was required. 1D and 2D NMR experiments have been used to confirm the ring closure as well as conformations by establishing through-bond and through-space connectivities.

The modifications in the sugar moieties of the nucleotides have shown their enormous potential for their efficiency in the sequence specific gene silencing and specific target binding. The incorporation of 1’,2’-constrained sugar moieties in nucleotides showed mixed results in their antisense and target affinity. 2’,4’-Constrained sugar moieties have shown better target affinities. Nevertheless, it was found that the nature of the nucleobase also plays a role in contributing conformational stability in addition to the sugar modification. The rigidity caused by the conformationally constrained nucleotides steers the structural and conformational preorganization as well as molecular recognition, for example the interaction with RNAse H in the antisense pathway. Furthermore, it can tune the dynamics of the oligonucleo-
tides. We have investigated the molecular structures and dynamics of 1’,2’– and 2’,4’–sugar constrained nucleosides in order to understand how the chemical nature of the conformational constraint of the sugar moiety influences the duplex structure and reactivity.

Papers I and II are comprehensive studies of the synthesis, structural analysis, and biochemical properties of conformationally constrained Aza-ENA and carbocyclic analogues. Some key points in the characterization of the modified nucleosides are further discussed below. The determination of the structures and the molecular dynamics study of the 1’,2’– and 2’,4’–sugar constrained nucleosides is discussed in Paper III.

2.1.1 Characterization of modified nucleosides

2.1.1.1 Aza-ENA-T

Two diastereomers, a major and a minor, were found to be the end products of the ring closure reaction in the synthesis of Aza-ENA-T (Paper I), and they could be separated. In order to identify the two isomers, a number of 1D NMR (1H, 13C, 1D NOE, 1D TOCSY) experiments, including single and double 1H homonuclear decoupling methods, and 2D NMR experiments (DQF-COSY, HMQC, and HMBC) have been performed. The multiplicity patterns of H7′ in 1H spectra have allowed the assignment of H7′e′ and H7′a′ in both major (8 lines, ddd) and minor (12 lines, dtd) diastereomers. It could be concluded, for the minor compound, that the multiplicity of 12 lines with three major coupling constants for H7′ was only possible if NH is in axial position. The chair conformations in both major and minor isomers have been confirmed by 1D NOE spectra by observing NOE correlations between H7′e′ and H1′. Furthermore, the HMBC spectra (Figure 9) not only confirmed the ring closure but also provided information related to possible dihedral angles φ[H7′e′-C7′-C2′-H2′] in the two isomers. The fact that the H7′e′ only have correlations with the C2′ confirmed that the dihedral angle φ[H7′e′-C7′-C2′-H2′] is closer to 180°, where the 3JHC is at maximum, and φ[H7′a′-C7′-C2′-H2′] is closer to 90°.

The minor isomer was found to have converted almost completely into the major isomer (>99%), in pyridine-d5 solution, irreversibly in 33h at 298K, and no reverse isomerism was observed starting from the major isomer under the experimental conditions. We have determined the rate constants at different temperatures under unimolecular first order rate kinetics by using 1H NMR spectroscopy, and subsequently E_a was determined to be 23.4 kcal mol⁻¹, and the frequency of collisions factor, A, to be 1.190 × 10¹⁴ s⁻¹. The free energy of activation, ΔG‡, was found to be 23.4 kcal mol⁻¹ at 298K in pyridine-d5 solution. In CDCl3 solution, the two isomers reach an equilibrium very slowly (in 30 days) with equilibrium constant K_c = 0.67, and
the $\Delta G^\ddagger$ was found to be 25.4 kcal mol$^{-1}$. This study suggested that the conversion is base catalyzed.

![Figure 9. Expansion of HMBC spectra of major and minor Aza-ENA diastereomers that confirm the ring closure. Observed correlations are indicated by double arrows. Reproduced from Varghese, O.P. et. al. J. Am. Chem. Soc. 2006, 128, 15173. Copyright ©2006 American Chemical Society.](image)

**2.1.1.2 Carbocyclic ENA and LNA**

For the five membered cyclized compound, two inseparable stereoisomers were found (in 7:3 ratio) as the end products in the free radical based ring closure reaction. The ring closure in the cyclization reaction for the mixture of both major and minor isomers could be confirmed by HMBC experiments (*Figure 10*). The complete resonance assignment for major and minor isomers was based on 1D and 2D NMR experiments. Stereochemistry including configuration at C7′ could be assigned on the basis of 1D NOE experiments. The major compound was confirmed to be the one having $R$ configuration at C7′ whereas the minor compound had $S$ configuration at C7′.
The analysis of NMR spectra for the 6-membered cyclized compound showed that it is indeed a single compound unlike the 5-membered one. The complete characterization was performed for the 6-membered cyclized compound and it was also confirmed to be the chair conformation having R configuration the C7′.
2.1.2 Determination of structure and dynamics of 1′,2′- and 2′,4′-constrained nucleosides

The reported experimental NMR coupling constants of conformationally constrained nucleosides containing T as the nucleobase have been used in the optimization of the group electronegativity parameter of the Haasnoot-de Leeuw-Altona\textsuperscript{159} generalized Karplus equation for the endocyclic nitrogen atom in the azetidine, Aza-ENA, and 2′-amino-LNA compounds as well as for the oxygen atom of the oxetane, ENA, and LNA-type nucleosides. The $^3J_{	ext{HH}}$ coupling constants and the corresponding torsion angle information obtained by \textit{ab initio} calculations have been used in order to obtain the optimized Karplus parameters using a least-square fitting numerical grid procedure. In addition, experimental coupling constants were further rationalized using 6-31G**Hartee-Fock optimized \textit{ab initio} gas phase calculations\textsuperscript{160}. NMR constrained molecular dynamics (MD) simulations\textsuperscript{161}, in explicit solvent, were performed which included an initial simulated annealing step followed by MD. The resulting structures were then used in constraints-free MD simulation steps.

It is found that, depending upon whether the modification leads to a bicyclic 1′,2′-fused or a 2′,4′-fused system, fall into two distinct categories characterized by their respective internal dynamics of the glycosidic and the backbone torsions as well as by characteristic sugar conformation: (i) \textit{North-East} type, $P = 37° ± 27°$ ($\Phi_m = 25° ± 18°$) of the 1′,2′-fused systems, and (ii) pure \textit{North} type, $P = 19° ± 8°$ ($\Phi_m = 48° ± 4°$) for the 2′,4′-fused nucleosides (\textit{Figure 11}).

In conclusion, the formation of the 2′,4′-conformationally constrained sugar containing thymidine nucleosides could be confirmed on the basis of long range through bond correlations and the conformations were confirmed for the major and minor isomers on the basis of NOE experiments. The NMR-derived conformational analysis followed by MD simulations showed that the sugar modification at the 1′,2′- and 2′,4′ have driven sugar conformations into \textit{NE} and pure \textit{N}-types, respectively.
2.2 Effect of sugar modification on the stability of DNA duplex: an NMR study (paper IV)

A variety of sugar-modified nucleosides have been incorporated into antisense oligonucleotides (AON) and tested for their physicochemical and biological properties such as stability, target affinity, and cleavage by nucleases. Depending on the modification and their site(s) of incorporation, the properties of the oligonucleotides were found to be different. The type of modification in the sugar moiety drives the conformational equilibrium of the nucleosides as well.

The 1’,2’-oxetane locked nucleosides are conformationally constrained to a unique fixed North-East type sugar conformation. The incorporation of oxetane modified thymidine (locked T, Figure 12) was shown to drop the \( T_m \) when incorporated into the antisense oligo/RNA hybrid, by \(<5\text{-}6 \, ^\circ\text{C}\) per each modification\(^{162-164}\). However, the incorporation of the oxetane-cytidine (C) moiety into the AONs imparts only \(<3 \, ^\circ\text{C}\) loss in \( T_m \) per oxetane-modification, whereas no loss in \( T_m \) has been observed for the oxetane-adenosine (A) or guanosine (G) modified AONs\(^{162, 164}\). Nevertheless, the RNase H recruitment capabilities of the oxetane-modified antisense oligo RNA hybrid duplexes were found to be very similar to that of the native counterpart\(^{164}\).

It was also shown that a locked T protects five RNA residues down stream of the base pairing with the locked T\(^{164}\) from RNAse H degradation.

\begin{table}
\centering
\begin{tabular}{|c|c|c|c|c|c|}
\hline
Azetidine-T (A) & Oxetane-T (B) & Aza-ENA-T (C) & ENA-T (D) & 2’-amino LNA-T (E) & LNA-T (F) \\
\hline
![](image) & ![](image) & ![](image) & ![](image) & ![](image) & ![](image) \\
\hline
RMSd = 0.569 & RMSd = 0.527 & RMSd = 0.433 & RMSd = 0.462 & RMSd = 0.567 & RMSd = 0.628 \\
(0.194, 0.739) & (0.171, 0.887) & (0.077, 0.668) & (0.078, 0.648) & (0.109, 0.699) & (0.089, 0.625) \\
\hline
\end{tabular}
\caption{Conformations of the 1’,2’- (A and B) and 2’,4’-(C-F) conformationally locked nucleosides. Reproduced from O. Plashkevych et. al. \textit{J. Org. Chem.} \textbf{2007}, \textit{72}, 4216. Copyright ©2007 American Chemical Society. The total RMSd values for all heavy atoms, heavy atoms of sugar moiety, and heavy atoms of base are shown in black, red, and blue, respectively.}
The endonuclease susceptibility of oxetane-modified antisense-oligos was also significantly reduced compared to the native counterpart, and it was proportional to the number of the oxetane modified nucleotides per AON molecule: single modification gave two-fold protection to the cleavage and double and triple modification gave four-fold protection compared to that of the native phosphodiester oligonucleotide.

The early work on the chronocoulometric measurements has shown that the charge transport through the DNA base stack is perturbed only in the DNA-DNA duplex and not in the DNA-RNA duplex containing modified T in the DNA strand. Although this qualitatively shows that the introduction of a constrained North-type sugar perturbs the helical stacked conformation of a DNA duplex, the study elucidates neither the nature of the conformational disorder created by the constrained conformation of the sugar, nor how far the conformational disorder propagates along the helix from the constrained site.

In order to understand the effect of inclusion of the oxetane-constrained sugar modification (Figure 12A) in the nucleoside on the DNA structure, the three dimensional solution structures of the native and modified Dickerson–Drew dodecamer (Figure 12) have been determined by NMR spectroscopy, where the oxetane modification was incorporated into T7 position (Figure 12B) of the self complementary strands.

![Diagram](image)

**Figure 12.** Oxetane locked T nucleoside (a) and the self complementary Dickerson-Drew dodecamer (b), d-5′(CGCGAATTGCAGC)₂.

Moreover, the melting behavior of the dodecamer was investigated by UV (at different salt concentrations) and by NMR spectroscopy. In addition, the base-pair imino proton dynamics have been studied using temperature de-
pendant \(^1\)H NOESY and NOESY-ROESY experiments, line-width analysis, and relaxation time measurements.

The CD spectra of the Dickerson–Drew dodecamer revealed that both native and modified dodecamers are in the right-handed B-form of the DNA duplex structure. The melting behavior at different salt concentrations showed differences in the native and modified counterparts. UV spectroscopy and 1D NMR spectra displayed a drop in melting temperature by <10 °C per modification for the T7 oxetane modified duplex compared to its native counterpart. The three dimensional structures of both the native and the T7 oxetane modified duplexes, obtained by NMR spectroscopy, have subsequently been compared with the results of 2.4 ns MD simulations.

The one dimensional NMR melting profiles of imino protons (Figure 13) as well as the aromatic and methyl protons of the nucleobases (Figure 14) reveal that the introduction of the \( NE\)-type locked T at the T7 position considerably perturbs the conformation of the four central base pairs and destabilizes the core of the modified duplex, promoting melting to start simultaneously from the center as well as from the ends.

![Figure 13. NMR Melting profiles of the Dickerson-Drew dodecamers (modified and native). NH of T7 and T8 show that they melt at different temperatures whereas those of the native counterpart melt simultaneously.](image-url)

33
Figure 14. NMR melting profiles of the native (I) and modified (II) Dickerson-Drew dodecamers. H6 of T7 and 5-Me of T8 show different melting temperatures, indicating that the melting behavior of the two residues is different whereas the native compounds melt simultaneously.

The NMR solution structures have been determined for the native and modified dodecamers (Figure 15). In spite of the drop in $T_m$, all bases participate in the base pairing in the modified dodecamer. The sugar-phosphate backbones of both the native (I) and the T7 modified (II) duplexes were found to reside mainly in typical B-type DNA conformation, i.e., $\alpha$, $\beta$, $\gamma$, $\epsilon$, and $\zeta$ with an exception for the parameters related to the T7 residue. The other geometrical parameters of the duplex also showed that the incorporation of the modified nucleoside into duplex (II) leads to only a few local changes at the site of modification. The sugar of the oxetane modified T7 residue maintained its NE-type conformation, while all other residues retain their typical B-type DNA conformations (S-type sugar). The backbone torsion angles are locally perturbed at the site of the lock T incorporation. The $\beta$ torsion is displaced from its normally preferred trans (ap) conformation to gauche+ (+sc), balanced by a shift of the $\gamma$ torsion from gauche+ (+sc) to trans (ap). The $\alpha$, $\epsilon$, $\zeta$, and $\chi$ torsions of the modified T7 residue are also shifted towards values associated with A-type conformation. The stacking is found to be less perturbed.

The central base pairs, A7T6 and T7A6 were strongly stretched and staggered. Consequently, they had non-linear base pairs (Figure 16). Moreover there is an increased roll between the two central modified base pairs compared to the native counterpart, which is compensated by increased tilts of the neighboring base steps. The temperature dependent hydration study by using 2D NOESY-ROESY experiments, $T_1$ relaxation studies, and line width analysis have also demonstrated that the central T7·A6/A6·T7 base pairs of the modified Dickerson-Drew dodecamer have at least one order of magni-
tude higher water exchange rates (correlated to the base pair opening rate) than the corresponding base pairs in the native duplex.

**Figure 15.** The NMR-derived secondary structures for the (a) native and (b) modified dodecamer. The native and modified residues in the central four base pairs are shown in blue and red, respectively.

**Figure 16.** Superimpositions of the T7:A6 base pair of the native (red) and the modified (green) forms show the perturbed geometry due to the modified oxetane-T.

In conclusion, the incorporation of the oxetane-locked nucleobases has affected the local perturbations in the central base pairs where the modification is included. However, the remaining structure was found to be unchanged. The oxetane-locked sugar was in NE-type whereas the other sugars were found to be in S-type conformations. The melting behavior observed by the aromatic and methyl protons as well as the exchangeable protons reflects the increased dynamics in the base pairs with oxetane modification. This study shows that the drop of melting temperature of antisense oligonucleotides is a result of the local perturbations in the structure caused by incorporation of the sugar-constrained nucleoside.
2.3 Effect of the sugar modification on the physicochemical properties of the nucleotides (paper III & V)

The aim of the paper IV was to investigate the influence of the chemical nature of the C2’-substituents (or modification) in the sugar moiety on the electronic properties of the nucleobases in model nucleosides (such as monoethyl or bisethyl phosphates). The change of the nature of the constrained ring from 1’,2’-fused to the 2’,4’-fused system was shown to change the equilibrium conformations from NE-type to N-type. The dependence of the size of the constrained ring of the sugar moiety on the pKa of the nucleobase as well as on the 2’-substituent has been explored using 1’,2’-fused or a 2’,4’-fused systems of thymidine nucleosides has been investigated in paper V. The pKa’s of the nucleotides were measured at 25°C by NMR acid-base titrations.

Acid base behavior of nucleotides constitutes their physicochemical characteristics. It determines the charge, tautomerism, and the ability to donate or accept hydrogen, which is a key feature in hydrogen bonding. The pKa is the measure of the strength of acidity and basicity of a molecule. The value of pKa depends on a number of variables such as component substituent effects, type of solvent, ionic strength, and salt concentration.

The pKa values of all the native nitrogenous bases in DNA and RNA show that they are well below or above the physiological pH. For being protonated or deprotonated at physiological pH, the pKa of the nucleobases should be reached closer to pH=7.0. The pKa perturbation has important structural and functional implications in biomolecules such as DNA, RNA, and proteins. In nucleic acids, numerous folded-state perturbed pKa towards neutrality have been reported. A+H . C mismatches with pKas of 6.5 include leadzyme, U6 loop of the spliceosome, a duplex, a loop of ribosomal RNA with cytosine having a pKa of 6.4 as some examples. The pKas of the nucleobases in the single stranded DNA and RNA has also been found to be modulated in a sequence dependent manner. The perturbation of pKa can occur as a result of: (i) interactions of charged groups; (ii) differences in solvent accessibility of charged groups; (iii) interactions such as hydrogen bonding and stacking; (iv) folding.

Any modification in the sugar moiety of the nucleoside has an effect on the pseudoaromaticity of its nucleobase. It is already shown that the nucleobases in deoxyriibo nucleotides are more basic than those in the ribonucleotides because of the electron withdrawing effect of 2’-OH in the latter. These modifications are shown to have profound consequences on the hydrogen bonding ability of the nucleobases. The exact extent of base pairing in a duplex is tunable by a sequence specific modulation of the pseudoaromaticity of nucleobases because of the nearest neighbor stacking interactions. Therefore any modulation observed in the acidity or basicity of the nucleobase by the chemical nature of the C2’ of the sugar moiety overwhelmingly influ-
ences the strength of hydrogen bonding as well as the stacking in potential duplexes or triplexes.

A 5'-phosphate group is known to increase the pKₐ of the nucleobase. Compared to the 2'-OMe group, N3 of pyrimidines (C, T, and U) or N7 of the imidazole group of G (purine) were found to be more acidic (Figure 17 shows the pKₐ of oxetane and 2'-OMe analogues of G and C, for the complete pKₐ data for all compounds, see Figure 1, page 1677 and Table 1, page 1676 in Paper V) because of a prominent electron withdrawing effect. On the other hand, N1 of purines (G, Figure 17) were shown to experience a poorer electron withdrawing effect.

In fact the pKₐ of N1/ N3 of the nucleobase in the 1',2'-conformationally constrained nucleotides can be a result of interplay between three effects depending on the 2'-substituent and/or the constrained ring: (i) effect of 5'-phosphate group (ii) anomeric effect (iii) distance dependent electron withdrawing effect between the C2' substituent and the protonation/ deprotona-

![Figure 17](image_url)

**Figure 17.** Experimental pKₐ values for oxetane-locked (upper row) and 2'-OMe (lower row) G and C. Values in brackets represent the pKₐ for 5', 3'-bisethyl phosphate compounds (R₁=R₂=[PO₃]⁻Et) whereas those in black are for 3'-monoethyl phosphates (R₁=H, R₂=[PO₃]⁻Et). The effect of 5'-phosphate is clearly evidenced. Effect of oxetane modification is reflected from the N7 of purines or N3 of pyrimidine nucleotides. The effect on pKₐ of N1 (in pyrimidine part) of purines is poorer as a result of the interplay of three effects: (i) phosphate (ii) anomeric effect (iii) electron withdrawing effect.
tion site of the nucleobase. Azetidine constrained ring also showed the same trend. There is a linear correlation between \( pK_a \) of N3 of pyrimidine (T/C/U) or N7 of 9-guaninyl with the corresponding H2′ chemical shift in NMR spectra at the neutral pH (Figure 18). It reflects that the pseudoaromatic character of the nucleobases can be tuned depending upon the chemical nature of the 2′-substituent. This high correlation of H2′ chemical shifts with the \( pK_a \) of the constituent nucleobase clearly suggests that the comparison of the chemical shifts for H2′ in the sugar moiety in a series of 2-modified nucleosides or oligonucleotides can provide neat information (bypassing direct \( pK_a \) measurement) for interrogation of how the pseudoaromatic character of the genetic alphabet has been altered as a function of the electronegativity of the 2′-substituent.

![Figure 18](image)

*Figure 18.* Representative plots of \( pK_a \) of nucleobase vs NMR chemical shifts of H2′ of the sugar moiety of nucleotides containing (a) U (b) G showing a high correlation dictating the chemical nature of the molecule. For more details, see Paper V. (Chatterjee, S. et al., *Org. & Biomol. Chem.* 2006, 4, 1675) – Reproduced by permission of The Royal Society of Chemistry.

This modulation of the \( pK_a \) of the nucleobase by a 2′-substituent is a through-bond as well as through-space effect, which has been supported by *ab initio* determined \( pK_a \) estimation. Interestingly, experimental \( pK_a \)s of nucleobases from NMR titration and the calculated \( pK_a \)s (by *ab initio* calculations utilizing closed shell HF 6-31G** basis set) are linearly correlated.

In order to understand the mechanism behind the influence of the 2′-modification on the \( pK_a \) of the nucleobase, a detailed analysis on the basis of frontier molecular orbitals has been performed. We have assumed that the highest occupied molecular orbitals (HOMO) and lowest unoccupied molecular orbitals (LUMO) as well as the lower occupied and higher unoccupied molecular orbitals HOMO of the nucleobase and the LUMO of the hydroxonium ion should play a role in the protonation/ deprotonation of the nucleobases, as the HOMO of the hydroxonium ion is much lower in energy.
than the LUMO of the nucleobase. The difference of ground and protonated/deprotonated HOMO for the nucleobases (A/G/C/T/U) are well correlated with their experimental $pK_a$s in different 2'-substituted 3,5-bis-ethylphosphate analogs, suggesting that only the orbital energy of the HOMO can be successfully used to predict the modulation of the chemical reactivity of the nucleobase by the 2'-substituent. Despite the high correlation between the HOMO energy and $pK_a$, it is believed that the HOMO orbital itself appeared to give limited clues for the deprotonation/protonation mechanism, as it was found to be a typical $\pi$ orbital, being 100% localized on the nucleobase for all purines and 55-98% for the pyrimidine nucleosides (MO for ground and N3-deprotonated as well as N7 protonated Oxetane G is shown in Figure 19a, Figure 19b, and Figure 19c, respectively, and those for ground and protonated Oxetane C are shown in Figure 20a and Figure 20b, respectively).

<table>
<thead>
<tr>
<th>HOMO-1</th>
<th>HOMO</th>
<th>LUMO</th>
<th>LUMO+1</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Figure 19. Molecular orbital diagrams for oxetane G in (a) the ground state (b) N1 deprotonated (c) N7 protonated forms (reproduced from supporting information section of: Chatterjee, S. et. al., *Org. & Biomol. Chem.* 2006, 4, 1675) – Reproduced by permission of The Royal Society of Chemistry.*

Electrons in HOMO-1 are shown to give highest contributions to the substitution effect. It has also been demonstrated that $pK_a$ values of nucleobases in 3,5-bis-ethylphosphates are well correlated with the change in dipole moment for the respective nucleobases after protonation or deprotonation.

To understand how the chemical nature of the conformational constraint of the sugar moiety in ON/RNA (DNA) dictates the duplex structure and
reactivity, we have determined molecular structures and dynamics of the conformationally constrained 1’,2’-azetidine- and 1’,2’-oxetane-fused thymidines, as well as their 2’,4’-fused thymine (T) counterparts such as LNA-T, 2’-amino LNA-T, ENA-T, and aza-ENA-T by NMR spectroscopy, *ab initio* (HF/6-31G** and B3LYP/6-31++G**), and molecular dynamics simulations (2 ns in the explicit aqueous medium).

<table>
<thead>
<tr>
<th></th>
<th>HOMO-1</th>
<th>HOMO</th>
<th>LUMO</th>
<th>LUMO+1</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td><img src="image" alt="Molecular orbital diagram" /></td>
<td><img src="image" alt="Molecular orbital diagram" /></td>
<td><img src="image" alt="Molecular orbital diagram" /></td>
<td><img src="image" alt="Molecular orbital diagram" /></td>
</tr>
<tr>
<td>b</td>
<td><img src="image" alt="Molecular orbital diagram" /></td>
<td><img src="image" alt="Molecular orbital diagram" /></td>
<td><img src="image" alt="Molecular orbital diagram" /></td>
<td><img src="image" alt="Molecular orbital diagram" /></td>
</tr>
</tbody>
</table>

*Figure 20. Molecular orbital diagrams for oxetane C in (a) the ground state (b) N3 protonated forms (reproduced from supporting information section of: Chatterjee, S. et. al., *Org. & Biomol. Chem.* 2006, 4, 1675) – Reproduced by permission of The Royal Society of Chemistry.*

As shown earlier (chapter 2.1.2) the modified bicyclic 1’,2’-fused or a 2’,4’-fused system have two distinct categories of conformations (pure N-type or NE-type). Each group has different conformational hyperspace accessible, despite the overall similarity of the N-type conformational constraints imposed by the 1’,2’- or 2’,4’-linked modification. The comparison of pKₐ’s of the 1-thyminyl aglycon as well as that of endocyclic sugar-nitrogen obtained by theoretical and experimental measurements have shown that the nature of the sugar conformational constraints steer the physicochemical property (pKₐ) of the constituent 1-thyminyl moiety, which in turn can play a part in tuning the strength of hydrogen bonding in base pairing: the basicity of the N3 of T as well as that of the 2’-N substituent has been found to be increased with the increase of the ring size (*Figure 21*).
Figure 21. $pK_a$ values of of the 2'-N and the nucleobase in the (A) 2'-NH$_2$-T, (B) Azetidine-T, (C) 2'-Amino-LNA T, and (D) Aza-ENA-T nucleosides. Both $pK_a$ for 2'-N and the N3 of T, respectively, show the tendency of increasing basicity with the increase of the constrained ring size.

This study provides evidence for modulating the chemical properties of the nucleobase in the nucleotide: (i) modifying the C2'-substituent (ii) constrained sugar molecules (1',2'- or 2',4'-) (iii) size of the ring size. Hence, the modulations of the $pK_a$ of the nucleobases caused by modifying the sugar moiety can be used to tune the hydrogen bonding. This provides an efficient way for designing molecules to be used in gene silencing.
3. Effect of CH-π interactions on stability of the hetero-duplexes (Paper VI)

Nucleic acid duplexes can occur in DNA-DNA (DD), RNA-RNA (RR), DNA-RNA (DR), or RNA-DNA (RD) forms. The order of thermodynamic stability of these duplexes is sequence dependent. Heteroduplexes, DR and DD, are found in cells as a result of transcription\textsuperscript{171}, Okazaki fragments in replication\textsuperscript{172}, in reverse transcription in retroviruses\textsuperscript{173}, and in the down regulation of genes by inhibition of mRNA using antisense technology\textsuperscript{27-29}. The understanding of their stability is important for the design of better molecular targets with specific functions.

The $pK_a$ differences ($\Delta pK_a$) between the 3'-monoethyl- [(d/rN)pEt] or 3',5'-bis-ethylphosphates [Etp(d/rN)pEt] in both 2-deoxy (dN) and ribo (rN) (N = A/G/C/T/U) nucleotides, have been qualitatively used to understand the strength of base-pairing energies in different DNA–RNA (DR), RNA–DNA (RD), DNA–DNA (DD), and RNA–RNA (RR) duplexes. These monomeric units have been used as the model donors and acceptors (in which stacking is completely eliminated) mimicking those of the internucleotide monomer building blocks of a duplex. The strengths of base pairing on the basis of $\Delta pK_a$ values are shown in (Table 3). Fourteen pairs of isosequential hybrid duplexes\textsuperscript{174} with varying sequence and length have been analyzed to dissect the relative energetic contributions from the base pairing [$G^\circ_{bp}$] vis-a-vis stacking [$G^\circ_{stacking}$] to the total free energy of stabilization [$G^\circ_{37}$]. Our experimental data from $pK_a$ studies have earlier shown\textsuperscript{96} that RR duplexes are more stable than DD duplexes because of larger energy gain from base pairing in the former compared to the latter. The aim of this study was to under-

<table>
<thead>
<tr>
<th>Base Pair</th>
<th>DD $\Delta pK_a$</th>
<th>DD $\Delta G_{(pK_a)}$</th>
<th>RR $\Delta pK_a$</th>
<th>RR $\Delta G_{(pK_a)}$</th>
<th>DR $\Delta pK_a$</th>
<th>DR $\Delta G_{(pK_a)}$</th>
<th>RD $\Delta pK_a$</th>
<th>RD $\Delta G_{(pK_a)}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>A:T(U)</td>
<td>6.29</td>
<td>35.9</td>
<td>5.53</td>
<td>31.6</td>
<td>5.43</td>
<td>31.1</td>
<td>6.43</td>
<td>36.4</td>
</tr>
<tr>
<td>T(U):A</td>
<td>6.43</td>
<td>36.4</td>
<td>5.43</td>
<td>31.1</td>
<td>6.43</td>
<td>36.4</td>
<td>5.43</td>
<td>36.4</td>
</tr>
<tr>
<td>GC</td>
<td>5.24</td>
<td>29.9</td>
<td>5.04</td>
<td>28.9</td>
<td>5.35</td>
<td>30.6</td>
<td>4.94</td>
<td>28.2</td>
</tr>
<tr>
<td>CG</td>
<td>4.94</td>
<td>28.2</td>
<td>5.35</td>
<td>28.2</td>
<td>5.35</td>
<td>30.6</td>
<td>4.94</td>
<td>28.2</td>
</tr>
</tbody>
</table>
stand the differences in the intrinsic nature of the electrostatic forces that are responsible for the self-assembly of heteroduplexes compared to homoduplexes.

The relative energetic contribution from the base pairing and stacking in the relative free energy of stabilization ($\Delta\Delta G^o_{37}$ from $T_m$) of the DR and RD has been dissected by examining the $G^o_{37}$ as a result of variation of the number of T in the D strand in the DR duplex vis-à-vis that in the RD duplex. This has been only possible in the heteroduplexes (DR/ RD) because T can be placed only in the DNA strand and U can be placed in the RNA strand in order to examine the effect of excess of %T in the D strand.

The differences in the free energy of stackings [$G^o_{stacking}$]$_{DR-RD}$ play a more significant role than the differences in the free energy of base-pairing, [$G^o_{bp}$]$_{DR-RD}$, in the relative stabilization of the DR or RD duplexes. It is found that, as the excess %T$_{DR-RD}$ increases, the $\Delta$T$_{m[DR-RD]}$ increases and vice-versa. When [%T]$_{DR-RD}$ are nearing zero, [T$_m$]$_{DR-RD}$ are also seen to be very close to zero (Figure 22).

![Figure 22](image.png)

**Figure 22.** Plot of relative stability of DR over RD versus % Excess T for fourteen isosequential heteroduplexes. The number in subscript indicates the sequence length (Chatterjee, S. et. al., Org. & Biomol. Chem. 2005, 3, 3911) – Reproduced by permission of The Royal Society of Chemistry.

This means that the other stacking interactions in the pairs are relatively small. The total relative stabilizations, [$G^o_{37}$]$_{DR-RD}$ among the DR and RD duplexes studied, are wholly dependent on the differences in the number of 5-Me(T) stacking interactions with the nearest neighbors in the D strands of DR duplexes compared to those of the RD duplexes.

**Figure 23** shows the dependence of excess %T on the relative contribution from base pairing and stacking, respectively. The relative $G^o_{bp}$ decreases as the number of excess Ts in DR over RD increases. As the number of excess T increases in DR compared to isosequential RD, the relative stacking stabilization in DR increases because of the methyl-T effect [Me(T)–Me(T) + Me(T)–(N), N = A/G/C/T, interactions].
The above findings indicate that the intrastrand stacking interaction between the C5-methyl group of 1-thyminyl and the cloud of the neighboring nucleobase as well as Me(T)–Me(T) interactions plays the dominant role over other stacking interactions.

This study shows that the incorporation of T in a heteroduplex in the DNA or RNA strand strengthens the duplex. This will be useful, for example, for the design of antisense oligonucleotides for better stability of the duplex.
4. Conclusion and outlook

The structure of the end products, (i) 2'-N, 4'-C-ethylene-bridged thymidine (Aza-ENA-T), and (ii) five and six membered 2',4'-carbocyclic ribothymidines, have been confirmed by 1D and 2D NMR spectroscopy.

The effects of 1', 2'– sugar constrained nucleotides on the structure and stability of the Dickerson-Drew dodecamer have been investigated by high resolution NMR, UV, and CD spectroscopy. Although the overall helical structure was maintained, local structural perturbations that resulted in stretched and staggered T7A6 and A7T6 non-linear base pairing have been observed. The central four base pairs are vulnerable to the exposure of water, resulting in faster melting behavior to resemble the nature of the fraying ends.

The pKₐ values of the nucleobase of several sugar constrained nucleosides and nucleotides have been studied. It is found that the pKₐ of the nucleobase is modulated depending on the chemical nature of the 2'– substituents as well as the constraint of the pentafuranose moiety and the extent of constraint. The major factors involved in the tuning of the electronic properties of a nucleobase in nucleotides were: (i) phosphate group, (ii) competing stereoelectronic effects (iii) electron withdrawing effect of the 2'– substituents. The strength of hydrogen bonding is dictated by the difference in pKₐ between the donor and acceptor atoms (pKₐ match). In a nutshell, the pseudoaromaticity and the physico-chemical properties of a nucleobase can be tuned by changing the 2'– modification, which results in a variation of the H-bonding strength of the nucleobase. This may be a strategy for designing molecular targets for specific recognition.

The incorporation of the aza-ENA, 2',4'–five and six membered constrained carbocyclic sugar modified nucleosides into the antisense oligonucleotides have shown to increase the stability. Moreover the cleavage patterns towards the RNAse H mediated degradation show that the extent as well as the cleavage sites is different depending on the modified nucleoside. This suggests that the incorporation of these modified nucleosides into the antisense oligonucleotides may have influence on their structure. This can further be investigated by determination of the structure and dynamics by NMR.

The earlier studies on pKₐ of nucleobases as well as the inter-nucleotidic phosphate groups in single stranded DNA and RNA show that their electronic properties and chemical reactivity are different depending upon the
sequence context. Furthermore, the study on the adenine-rich hexameric single stranded DNA and RNA structures have earlier shown that they are pre-organized. This work could be extended by investigation of the effect of sequence specificity on the single stranded DNA or RNA structure.
This work has been done in the Department of Biochemistry and Organic Chemistry and the Department of Bioorganic Chemistry. I take this opportunity to thank all who helped and supported me.

First of all, I owe a great debt of gratitude to Associate Professor Adolf Gogoll and Professor Helena Grennberg for accepting and supporting me as a doctoral student in the group to complete my Ph.D. studies. Your help and critical comments for my thesis work are invaluable. I wish to thank the administration of Uppsala University to provide me a stable platform to defend my thesis. I would like to thank the department head Prof. Lars Baltzer, and all the members of the Department of Biochemistry and Organic Chemistry, and the members of the Faculty of Science and Technology for their sincere help for achieving my goal. I wish to express my sincere gratitude to Prof. Jyoti Chattopadhyaya for being a mentor, providing excellent laboratory facilities with high field NMR, and financing my research. I wish to extend my heartiest thanks to Dr. Oleksandr Plashkevych for being an assistant supervisor and for assisting me learning molecular dynamics and computational work.

I gratefully acknowledge the support and supervision of late Associate Professor Premila Perera-Ivarsson, Prof. Lars Bohlin and all the present and past members of the Division of Pharmacognosy, Department of Medicinal Chemistry, Uppsala University.

I would like to thank Dr. P. I. Pradeep Kumar, Dr. Anders Földesi, Dr. Thirumari Mariappan, Dr. Petr Naus, Dr. Naresh Badgujar, Dr. Mounir Andaloussi, Pradeep Cheruku, Dimitri Kulik, Ayan Samanta, Hu Wenbin, and Zhou Chuanzheng for providing me a friendly atmosphere in the lab.

My special thanks goes to:

Subhrangsu Chatterjee for being my colleague and co-worker, having fruitful discussions in our research, encouraging me all the time, and providing me a very cooperative and friendly atmosphere inside and outside the lab in everyday life during my whole Ph.D research work.
Dr. Sandipta Acharya and Dr. Parag Acharya for their support, discussions and initial support in the NMR lab.

Dr. Johan Isaksson for his help for first lessons in molecular modeling by using NMR.

Puneet Srivastava, for being a co-worker and having discussions on scientific and other different subjects, Oommen P. Varghese, Jharna Barman, Mohitosh Maiti, Dimitri Honcharenko, Dr. Malgozarta Wenska for being co-workers and having friendly atmosphere in the lab during my research.

Jaana Evander and Eva Pylvänen for their administrative and secretarial support. I wish to thank all the members of staff in the BMC library and BMC administration.

I wish to pay my gratitude to my teachers in the school and university days. I am indebted for Prof. Ananda Amarasekara for introducing me to the long run of my research career. I would like to thank all the academic and non-academic staff members of the University of Colombo, Sri Lanka especially to Dr. Sujatha Hewage and the academic and non-academic staff members of the Department of Chemistry.

It is my pleasure to thank Dr. Wimal Uabhayasekara, Kumari Uabhayasekara, Hemanthi Ratnayake, Dr. Jagathpriya Weerasena, Sanjeewani Sooriyarachchi, Dr. Mahinda Kommalage, Gemini Manjaree, Dr. Dinesh Fernando, Hasanthi Fernando, Prof. Jayantha Welihinda, Lalith Walisadeera, Samanthi Priyadashanee, Kalyani Munasinghe, Malik Silva, Dr. Gamini Piyadasa and their families. Your friendship, invaluable assistance and encouragement made me comfortable in Uppsala. I wish to thank all the other Sri Lankan friends inside and outside Sweden including Dr. Wilfred Wijesinghe, Lakshman Perera, Samantha Tennakoon, J. K. Nikawela, Suraj Perera, Gomika Udugamasooriya, Chandrika Palamakumbura as well as non-Sri Lankan friends for all the encouragements and support.

I wish to thank Dr. Suparna Sanyal, Dr. Biplab Sanyal, Dr. Jayadip Ghosh, Gopal Datta, Prof. Ashok Swain, Dr. Ranjula Bali Swain, Dr. Santanu Dasgupta, Debashree Ghosal for their friendship.

I would like to thank Per Löwdin, Mattias Wiberg, Anneleen Kool and Jenny Karlsson for their humble support.

I am extremely grateful to my mother, father, late brother, and sisters, in laws, and other relatives for caring me all the time and being patient for my
absence in the important events in the family. I am greatly indebted to my aunt and late uncle for all your care, kindness and love.

Last but not the least I do not have words to thank my beloved wife, Ayesha, for your dedication and patience throughout my research studies. I could not achieve this goal without your help by taking care of our family, especially during my late hours of work. My son Kanishka and daughter Thilini your dedication and patience made my life easier. You were my only happiness after exhausting long days.
Nukleinsyrorna utgör en grupp av polymerer som förekommer i biologiska system. De förekommer i två olika huvudformer: Deoxyribonukleinsyra (DNA) och ribonukleinsyra (RNA). Monomererna kallas för nukleotider. Dessa är fosforsyraestrar av nukleosider, vilka i sin tur är uppbyggda av en sockerenhet (monosackarid) och en organisk bas. I monomererna av DNA och RNA förekommer vardera fyra organiska baser, tre av dessa förekommer i både DNA och RNA. Storleken av nukleinsyrorna varierar mellan 80 nukleotider i tRNA och flera miljoner nukleotider i kromosomernas DNA.


<table>
<thead>
<tr>
<th>A-DNA</th>
<th>B-DNA</th>
<th>Z-DNA</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image" alt="A-DNA" /></td>
<td><img src="image" alt="B-DNA" /></td>
<td><img src="image" alt="Z-DNA" /></td>
</tr>
</tbody>
</table>

*Bild 1. Olika strukturella former av nukleinsyror som bildar dubbelhelix.*

Även strukturellt annorlunda nukleotider förekommer naturligt i celler, och har biologisk funktion. Kemiskt modifierade, dvs artificiella nukleotider och nukleosider, används för att undertrycka vissa geners funktion (som antisense, antigen, och RNAi), och de har intressanta tillämpningar som molekylära testsubstanser för att undersöka strukturer och reaktionsmekanismer där nuk-
Leinsyror är involverade. Som exempel kan tas antisenseteknologin, där viss proteinskyntes förhindras genom att en artificiell oligonukleotid binds till mRNA (Bild 2).

Bild 2. Den genetiska koden lagras i DNA. Koden översförs till mRNA, och översätts därefter till proteiner, som i sin tur ger cellers funktion. Om en antisensenukleotid binds till mRNA kan inte protein bildas. Detta kan användas som läkemedel som fungerar på gennivå vilket kan vara mer effektivt än traditionella läkemedel som istället interagerar med proteiner.

En av de vanligaste metoderna för att ta fram modifierade nukleotider är kemisk modifiering av sockerdelen. Flera möjligheter förekommer: Modifiering vid C2'-positionen (t ex 2'-OMe), eller konformationsrestriktion vid C1' och C2' eller C2' och 4'. Det är önskvärt att modifieringarna leder till en rad av effekter: (i) varierad affinitet i samband med molekylära interaktioner, (ii) specificitet, (iii) stabilitet gentemot nedbrytande endo- och exonukleaser och (iv) förmåga att ansamlas i rätt typ av vävnad, för att den farmakologiska effekten ska bli bättre.

Denna avhandling beskriver i första hand strukturutredning, konformationsanalys och bestämning av fysikalisk-kemiska egenskaper hos nukleotider och oligonukleotider mha magnetisk kärnresonansspektroskopi (NMR).

Publikationerna I och II beskriver syntes av C2' och C4' modifierade sockerarter, deras inbyggnad i oligonukleotider, samt en analys av deras fysiologiska effekter. Det visas att dessa modifieringar resulterar, utöver ändrade kemiska egenskaper av sockret, även i ändrade kemiska egenskaper av nukleotiderna och molekylernas dynamiska egenskaper. I publikation III beskrivs molekylära strukturer samt dynamiska egenskaper av några nukleosider med socker modifierade vid C1', C2' eller C2', C4' baserat på resultat från NMR-spektroskopi, ab initio beräkningar, och molekyldynamiksimuleringar.

Inbyggnad av modifierade nukleotider i antisenseoligonukleotider (publikation IV) visade sig ge upphov till förändrad duplexstabilitet och även för-
ändrad interaktion med det RNAse som degraderar mRNA. För att förstå hur sockerenhetens flexibilitet påverkar DNAs duplexstruktur har nukleotider med oxetan-modifierade sockerenhet byggts in i en dodekaoligonukleotid med antisenseegenskaper. Strukturen hos oligonukleotiden har bestämts med högupplöst NMR-spektroskopi. Temperaturstabilitet och dynamik har även det studerats med NMR-tekniker, ssk NOESY och NOESY-ROESY-experiment. Den modifierade nukleotiden visade sig ha en annorlunda vätebindningsgeometri (Bild 3) som leder till en störning i duplexstrukturen jämfört med en oligonukleotid med naturlig nukleotid.

![Bild 3. En DNA dodekamer med modifierad T-nukleotid (grön) har inte samma bindningsgeometri som en med omodifierad T-nukleoosid (röd). Även hydatiseringsmönstret påverkas (ej visat i bilden).](image)


En viktig parameter i antisensteknologin är stabiliteten hos heteroduplexstrukturer av DNA och RNA, då detta påverkar den molekylära ingenkännningen hos nukleinsyror.

DNA har en metylgrupp i thymidinets 5-position som bidrar till stabiliteten hos DNA-RNA hybrider genom CH-π interaktioner. Detta diskuteras i publikation VI.
7. References


59


A doctoral dissertation from the Faculty of Science and Technology, Uppsala University, is usually a summary of a number of papers. A few copies of the complete dissertation are kept at major Swedish research libraries, while the summary alone is distributed internationally through the series Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology. (Prior to January, 2005, the series was published under the title “Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology”.)