Studies on the Non-covalent Interactions (Stereoelectronics, Stacking and Hydrogen Bonding) in the Self-assembly of DNA and RNA

BY

PARAG ACHARYA

ACTA UNIVERSITATIS UPSALIENSIS
UPPSALA 2003
Dissertation presented at Uppsala University to be publicly examined in B7:113a, Floor 1, Biomedical Centre, Uppsala, Friday, December 19, 2003 at 13:00 for the degree of Doctor of Philosophy. The examination will be conducted in English.

Abstract

This thesis is based on ten publications (Papers I-X). The phosphodiesters backbone makes DNA or RNA to behave as polyelectrolyte, the pentose sugar gives the flexibility, and the glycones promote the self-assembly or the ligand-binding process. The hydrogen bonding, stacking, stereoelectronics and hydration are few of the important non-covalent forces dictating the self-assembly of DNA/RNA. The pH-dependent thermodynamics clearly show (Papers I and II) that a change of the electronic character of aglycone modulates the conformation of the sugar moiety by the tunable interplay of stereoelectronic anomeric and gauche effects, which are further transmitted to steer the sugar-phosphate backbone conformation in a cooperative manner. 3'-anthraniloyl adenosine (a mimic of 3'-terminal CC3CHC of the aminoacyl-tRNA) binds to EF-Tu*GTP in preference over 2'-anthraniloyl adenosine, thereby showing (Paper III) that the 2'-endo sugar conformation is a more suitable mimic of the transition state geometry than the 3'-endo conformation in discriminating between correctly and incorrectly charged aminoacyl-tRNA by EF-Tu during protein synthesis. The presence of 2'-OH in RNA distinguishes it from DNA both functionally as well as structurally. This work (Paper IV) provides straightforward NMR evidence to show that the 2'-OH is intramolecularly hydrogen bonded with the vicinal 3'-oxygen, and the exposure of the 3'-phosphate of the ribonucleotides to the bulk water determines the availability of the bound water around the vicinal 2'-OH, which then can play various functional role through inter- or intramolecular interactions. The pH-dependent 1H NMR study with nicotinamide derivatives demonstrates (Paper V) that the cascade of intramolecular cation (pyridinium)-π(phenyl)-CH=CH(methyl) interaction in edge-to-face geometry is responsible for perturbing the pKa of the pyridine-nitrogen as well as for the modulation of the aromatic character of the neighboring phenyl moiety, which is also supported by the T1 relaxation studies and ab initio calculations. It has been found (Papers VI-IX) that the variable intramolecular electrostatic interaction between electronically coupled nearest neighbor nucleobases (steered by their respective microenvironments) can modulate their respective pseudoaromatic characters. The net result of this pseudoaromatic cross-modulation is the creation of a unique set of aglycones in an oligo or polynucleotide, whose physico-chemical properties are completely dependent upon the propensity and geometry of the nearest neighbor interactions (extended genetic code). The propagation of the interplay of these electrostatic interactions across the hexameric ssDNA chain is considerably less favoured (effectively up to the fourth nucleobase) compared to that of the isosequential ssRNA (up to the sixth nucleobase). The dissection of the relative strength of basepairing and stacking in a duplex shows that stability of DNA-DNA duplex weakens over the corresponding RNA-RNA duplexes with the increasing content of A-T/U base pairs, while the strength of stacking of A-T rich DNA-DNA duplex increases in comparison with A-U rich sequence in RNA-RNA duplexes (Paper X).

Keywords: nucleic acids, NMR, Stereoelectronic effects, Hydrogen Bonding, Stacking

Parag Acharya, Department of Bioorganic Chemistry, Box 581, Uppsala University, SE-75123 Uppsala, Sweden

© Parag Acharya 2003

ISSN 1104-232X
ISBN 91-554-5821-1
urn:nbn:se:uu:diva-3825 (http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-3825)
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
## Contents

1. Self-assembly of nucleic acid  
   1.1 Conformation of nucleic acids  
   1.2 Forces responsible for the self-assembly  

2. Stereoelectronic effects in nucleosides and nucleotides  
   2.1 The pseudorotation concept for pentofuranose  
      2.1.1 Two-state North (N) ⇆ South (S) pseudorotational equilibrium  
      2.1.2 Pseudorotational energy barrier for pentofuranose  
   2.2 Types of stereoelectronic forces and their interplay  
      2.2.1 Tunable anomeric effect  
      2.2.2 Tunable gauche effects  
      2.2.3 Configuration dependent gauche effect  
      2.2.4 The interplay of competing anomeric and gauche effects  
   2.3 Biological implications of stereoelectronic effects in nucleic acids  

3. Present Work (Papers I - III)  
   3.1 Aglycone-sugar-phosphate conformational cooperativity  
      (Papers I and II)  
      3.1.1 pH-dependent shift of N ⇆ S pseudorotational equilibrium  
      and ε ⇆ ε’ equilibrium in nucleotides  
   3.1.2 The cooperative shift of the (N,ε) ⇆ (S,ε') in nucleotides  
   3.1.3 Tunability of aglycones and transmission of the electronic character  
   3.1.4 Mechanistic overview of RNA as molecular wire for conformational transmission (Paper I)
3.2 *Strength of gauche effect dictates the EF-Tu recognition of aminoacyl-tRNA: studies based on the mimicking model* (Paper III) 16

3.3 Implications 17

4. The contribution of H-bonding in biomolecular interactions 18

4.1 *Nature of hydrogen bonding* 18
  4.1.1 Hydrogen bonding in nucleic acids 19
  4.1.2 Hydrogen bonding in peptides and carbohydrates 19

4.2 *Hydrogen bonding by 2’-hydroxyl group in RNA* 19
  4.2.1 Importance of 2’-hydroxyl group in RNA 20
  4.2.2 Hydrogen bonding and hydration of 2’-hydroxyl group in RNA 20

5. Present Work (Paper IV) 21

5.1 *Hydrogen bonding and hydration of 2’-hydroxyl group in nucleosides and nucleotide* 21
  5.1.1 Geometry of intramolecular hydrogen bonding 21
  5.1.2 Thermodynamics of inter- and intramolecular hydrogen bond 22
  5.1.3 Nature of hydration around the 2’-hydroxyl group 22

5.2 Implications 23

6. Aromatic interactions and stacking 23

6.1 *Aromatic interactions* 23
  6.1.1 Aromatic interactions in non-biological model system 24
  6.1.2 Aromatic interactions in biological system 26

6.2 *Base-base stacking interaction in nucleic acids* 26

6.3 *pK_a perturbation in nucleic acids and protein folding* 27
  6.3.1 pK_a perturbation in nucleic acids 27
  6.3.2 pK_a perturbation in proteins 28

6.4 *Biological importance of single starnded nucleic acids* 29

7. Present Work (Papers V - X) 31

7.1 *Tandem nearest-neighbor aromatic interactions in nicotinamide derivative* (Paper V) 31
7.1.1 Electrostatic interactions mediated $pK_a$ perturbation of pyridinyl group 31
7.1.2 Nearest-neighbor interaction between pyridinyl and phenyl groups 32
7.1.3 Cascade of pyridinium-phenyl-methyl cross-talk 32

7.2 Implications 32

7.3 Cross-modulation of physicochemical character of nucleobases in single stranded nucleic acids (Papers VI - X) 33
7.3.1 The pH-dependent titration of dinucleotides and other ssRNAs and ssDNAs (Papers VI - IX) 33
7.3.2 Sequence specific nearest-neighbor interaction and thermodynamics of the offset stacking (Papers VI and IX) 35
7.3.3 The $pK_a$ shift of 9-guaninyl as a result of electrostatic effect due to 3'- and 5'-phosphate versus nearest-neighbor nucleobase (Papers VI and VII) 36
7.3.4 Variation of $pK_a$ of 9-guaninyl among different marker protons across the single strand (Papers VII - IX) 36
7.3.5 Propagation of electrostatic interplay across the single strand (Papers VIII and IX) 37
7.3.6 ssDNA is better stacked than that of ssRNA (Papers VIII and IX) 37
7.3.7 Tentative stacking geometry in single stranded nucleic acid from NMR constrained MD simulation: a qualitative approach (Paper IX) 38
7.3.8 Relative dissection of stacking vis-à-vis basepairing from the $pK_a$ calculations of the model mononucleotides (Paper X) 38

7.4 Implications 39

8. Acknowledgements 41

9. References 43
## Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>AE</td>
<td>Anomeric effect</td>
</tr>
<tr>
<td>B</td>
<td>Nucleobase</td>
</tr>
<tr>
<td>D</td>
<td>Deprotonated state</td>
</tr>
<tr>
<td>DD duplex</td>
<td>DNA-DNA duplex</td>
</tr>
<tr>
<td>DNA</td>
<td>DeoxyriboNucleic Acid</td>
</tr>
<tr>
<td>DMSO</td>
<td>Dimethyl Sulphoxide</td>
</tr>
<tr>
<td>E</td>
<td>Envelope</td>
</tr>
<tr>
<td>EF-Tu</td>
<td>Elongation Factor Tu</td>
</tr>
<tr>
<td>GE</td>
<td><em>Gauche</em> effect</td>
</tr>
<tr>
<td>H-bond</td>
<td>Hydrogen bond</td>
</tr>
<tr>
<td>N</td>
<td>Neutral state</td>
</tr>
<tr>
<td>N-type</td>
<td>North type</td>
</tr>
<tr>
<td>NMR</td>
<td>Nuclear Magnetic Resonance</td>
</tr>
<tr>
<td>MM</td>
<td>Molecular Mechanics</td>
</tr>
<tr>
<td>MP2</td>
<td>Møller Plasset (2nd order) basis set</td>
</tr>
<tr>
<td>mRNA</td>
<td>Messenger RNA</td>
</tr>
<tr>
<td>P</td>
<td>Protonated state</td>
</tr>
<tr>
<td>PNA</td>
<td>Peptide Nucleic Acid</td>
</tr>
<tr>
<td>R</td>
<td>Pearson Correlation Coefficient</td>
</tr>
<tr>
<td>RNA</td>
<td>Ribonucleic Acid</td>
</tr>
<tr>
<td>ROESY</td>
<td>Rotating-frame nOe spectroscopy</td>
</tr>
<tr>
<td>RR duplex</td>
<td>RNA-RNA duplex</td>
</tr>
<tr>
<td>S-type</td>
<td>South type</td>
</tr>
<tr>
<td>ss</td>
<td>Single stranded</td>
</tr>
<tr>
<td>T</td>
<td>Twist</td>
</tr>
<tr>
<td>tRNA</td>
<td>Transfer RNA</td>
</tr>
<tr>
<td>$\Delta G$</td>
<td>Free energy</td>
</tr>
<tr>
<td>$\Delta H$</td>
<td>Enthalpy</td>
</tr>
</tbody>
</table>
1. The self-assembly of nucleic acids

1.1 Conformation of nucleic acids

The various functionalities and conformation of nucleic acid dictate its biological recognition, interaction and activity. Endowed with unique capabilities, such as the storage of the genetic information, induction of cellular differentiation as well as splicing, self-cleavage and catalysis, DNA and RNA are nucleic acid polymers built up of monomeric nucleosides.

Scheme 1. Panel (A) shows the endocyclic ($\nu_o - \nu_4$) torsions (see section 2.1.1 for details) of pentofuranose and sugar-phosphate backbone torsions ($\alpha$, $\beta$, $\gamma$, $\delta$, $\varepsilon$ and $\zeta$; Newman projections of gauche$^+$, trans and gauche$^-$ orientations for $\beta$, $\gamma$ and $\varepsilon$ are also shown). Panel (B) shows the constituent nucleobases (with atom numbering) for DNA and RNA along with ionization site (with small arrow) and the representative p$K_a$ values for nucleosides.\textsuperscript{1}

which are covalently linked through $3'\rightarrow 5'$-phosphodiester linkages.\textsuperscript{1,2} Nucleosides consist of a D-pentofuranose sugar and a heterocyclic nucleobase [adenin-9-yl (A), guanin-9-yl (G), cytosin-1-yl (C), thymin-1-yl (T), uracil-1-yl (U), A/C/G/T for DNA with $\beta$-$\text{D}$-2$'$-deoxyribosyl and A/C/G/U for RNA with $\beta$-$\text{D}$-ribofuranosyl sugar, Scheme 1].\textsuperscript{1,2} The conformation along the sugar-phosphate backbone in nucleotides can be fully defined\textsuperscript{1,2} by the torsion angles $\alpha$, $\beta$, $\gamma$, $\delta$, $\varepsilon$ and $\zeta$ (Scheme 1A). The conformational change across these torsions induced by various non-covalent forces (section 1.2) ultimately dictates the self-assembly of nucleic acids.
1.2 Forces responsible for the self-assembly

The phosphodiester moieties at the backbone makes DNA or RNA to behave as polyelectrolyte, the pentose sugar gives the flexibility, and the aglycones help in the self-assembly or the ligand-binding process. It is still unclear whether the sugar moiety drives the phosphate backbone geometry or the phosphate drives the sugar conformation. The non-covalent forces\(^1,2\) which dictate the self-assembly of nucleic acids can be attributed to the following: (i) Strong intermolecular H-bonds between complementary nucleobases in the opposite strands; (ii) Intra and intermolecular hydrogen-bonding with the 2'-OH group in RNA (sections 4.2 and 5.1); (iii) Intramolecular base-base stacking interactions\(^1,2\) (electrostatic interactions and hydrophobic forces between adjacent base pairs and across the nucleotide chain), (iv) hydration\(^1,2,4\) in the minor and major grooves, and (v) stereoelectronic gauche\(^7-9,57-59,61-63,67,69,73-74\) and anomeric\(^7-9,36,57-59,68-70,72,74,75\) effects within the sugar-phosphate backbone.

The aim of the ongoing studies has been to explore how a change of the local environment orchestrates the interdependent physico-chemical behavior of the aglycone, sugar and phosphate in nucleic acids in a coordinated manner. It has thus emerged\(^62-75\) from the studies performed in this laboratory that the net local conformational changes result from an interplay of stereoelectronics forces (section 2), whereas the electrostatic forces between the nearest neighbor nucleobases (section 7) modulate each others pseudoaromatic character, thereby altering their chemical reactivity depending upon the sequence context.

In order to quantify these weak non-covalent forces responsible for the self-assembly of nucleic acids, we have assessed the thermodynamics of various conformational equilibria along the sugar-phosphate backbone torsions to understand the energetic preferences of conformational states. Similarly, we have employed the pH-dependent chemical shifts, \(pK_a\) measurements or the NMR relaxation process as tools to understand pH-dependent stacking ⇋ destacking equilibrium among the nearest neighbor nucleobases (section 7.3) to evaluate the nature of transmission of electrostatic interactions through the single-stranded DNA and RNA, which are indeed the active intermediates for replication, transcription or translation (in conjunction with various specific proteins).

2. Stereoelectronic effects in nucleosides and nucleotides

2.1 The pseudorotation concept for pentofuranose

The pseudorotation concept has been introduced by Kilpatrick et al.\(^10\) in order to describe the continuous interconversions between the puckered forms of
the cyclopentane ring. A barrier to planarity of cyclopentane of 22 kJ mol\(^{-1}\) has been reported.\(^{11}\) Thus, the cyclopentane ring relieves its strains by pseudorotation, which would be induced by 120° bond angles and eclipsed methylene groups if it would adopt a planar geometry.\(^{12}\)

The puckered geometry of the pentofuranose ring in nucleic acid derivatives can be described\(^{13-16}\) by two parameters: the phase angle of pseudorotation (\(P\), showing which part of the ring is mostly puckered) and maximum puckering amplitude (\(\Psi_m\), showing the largest deviation of the endocyclic torsions from zero).

2.1.1 Two-state North (N) ⇋ South (S) pseudorotational equilibrium

Both the results of NMR studies\(^{17-22}\) (two distinctly identifiable and dynamically interconverting N- and S-type conformations have been observed in some B ⇋ Z DNA\(^{17,18}\), A ⇋ Z RNA\(^{19,20}\) or A-form ⇋ B-form lariat RNA\(^{21,22}\) transitions) as well as statistical analysis\(^{23}\) of the distribution of \(P\) values of the crystal structures of nucleos(t)ides (Scheme 2) by de Leeuw et al suggest that the conformation of the pentofuranose can be adequately described by a two-state North (N, C3'-endo, C2'-exo) ⇋ (S, C2'-endo, C3'-exo) equilibrium model, since no third state is found yet. Only a few E-type pseudorotamers\(^{24}\) and no W-type conformers were found among these crystal structures. W-type conformers are energetically disfavoured owing to the pseudoaxial orientation of both the nucleobase and the 5'-CH\(_2\)OH group as well as the eclipsed C2' and C3' substituents (Scheme 2). Analogously, the energy destabilization of E-type conformations (compared to either N- or S-type conformation) can be attributed to the eclipsed orientation of the C2' and C3' substituents.

\(P\) and \(\Psi_m\) are related to the five endocyclic torsion angles \(\nu_i\) (\(i = 0...4\)) (Eq 1):

\[\nu_i = \Psi_m \cos (P + 4\pi (i - 2) / 5)\]  

..... Eq 1
The ensemble of puckered forms of the pentofuranose in nucleosides is represented in the form of the pseudorotation cycle (Scheme 2). In β-D-nucleosides, the endocyclic torsions \( \nu_i \) (\( i = 0 \) to \( 4 \)) are defined as follows: \( \nu_0 [C4'-O4'-C1'-C2'] \), \( \nu_1 [O4'-C1'-C2'-C3'] \), \( \nu_2 [C1'-C2'-C3'-C4'] \), \( \nu_3 [C2'-C3'-C4'-O4'] \) and \( \nu_4 [C3'-C4'-O4'-C1'] \) (Scheme 1). In a study of 178 β-D-furanosides, the endocyclic torsion angles could be calculated using Eq 1 with an r.m.s. error of 0.4 - 0.9°.

Altuna’s model has also been used to describe the conformation of the puckered states of cyclopentane, the pyrrolidine ring in L-proline and in its derivatives, the ring D in steroids, and other five-membered rings.

### 2.1.2 Pseudorotational energy barrier for pentofuranose

For cyclohexane, the energy barrier for interconversions between two chair forms (Scheme 3) is ~45 kJ mol\(^{-1}\). Similarly, the energy for ring inversion between two chair forms for 2-methoxy-1,3-dimethylhexahydropyrimidine is ~37 ± 2 kJ mol\(^{-1}\) as calculated from \(^1\)H and \(^13\)C NMR measurement. The pseudorotational barrier for β-D-glucose is ~46 kJ mol\(^{-1}\).

Table 1 in ref. 7 showed the effect of substituent (X) at the anomeric position of a heterocyclic six-membered ring: \( \Delta G_{\text{ax/eq}} \) is ranging from 0.1 kJ mol\(^{-1}\) (for \( X = \text{CO}_2\text{Me} \)) – 9.6 kJ mol\(^{-1}\) (for \( X = \text{Br} \)). However, hexopyranoses are much less flexible (in NMR time scale) than pentofuranose (Scheme 3), as the latter having higher energy barrier compared to the former (\( \Delta E^\dagger_{\text{pyranose}} > \Delta E^\dagger_{\text{furanose}} \)). It has been shown that the polysaccharide elasticity is governed by the chair-boat transition of glucopyranose ring.
Theoretical calculations have been performed to estimate the energy barrier for pseudorotation of five and six membered rings. The height of barrier for pseudorotation of tetrahydrofuran is measured\textsuperscript{39} to be 2.9 kJ mol\textsuperscript{-1}. The \textit{ab initio} calculations\textsuperscript{40} showed the energy barrier for the pseudorotation of pyrrolidine as 2.5 kJ mol\textsuperscript{-1}. However, the theoretical calculations on 1-amino derivative of ribose, 2'-deoxyribose or 3'-deoxyribose\textsuperscript{41,42} and on nucleosides themselves\textsuperscript{43} have shown that the activation energy barrier for N- to S- interconversions is clearly greater in the W (\(\approx 24\) kJ mol\textsuperscript{-1} for 2 deoxyribofuranose, \(\approx 31\) kJ mol\textsuperscript{-1} for ribofuranose) than in the E region (\(\approx 7.5\) kJ mol\textsuperscript{-1} for 2'-deoxyribofuranose, \(\approx 16\) kJ mol\textsuperscript{-1} for ribofuranose). The pseudorotational barrier of furanose in both ribose and 2'-deoxyribose have been calculated to be \(\approx 2.5\) kJ mol\textsuperscript{-1} from consistent force field method\textsuperscript{44}, however, the classical energy calculations\textsuperscript{45} have found the pseudorotational barrier as \(\approx 16.7\) kJ mol\textsuperscript{-1} for ribopurines and 12.6 – 16.7 kJ mol\textsuperscript{-1} for ribopyrimidines. PCILO calculations showed\textsuperscript{46,47} that the height of the pseudorotation barrier for 2'-deoxyadenosine and 2'-deoxyuridine their ribo counterparts varies in the range 16.5 – 23.0 kJ mol\textsuperscript{-1}. Similarly, the height of the pseudorotation barrier for ribonucleosides and 2'-deoxynucleosides has been estimated\textsuperscript{48} to be equal to \(\approx 16.7\) and 10.5 kJ mol\textsuperscript{-1} respectively. The \textit{ab initio} (MP2/6-31G*)\textsuperscript{39} calculations with imidazole 2'-deoxynucleoside and ribonucleoside analogue as well as with natural 2'-deoxynucleosides\textsuperscript{50,51} (with geometrical parameters from crystal structures of nucleosides) have shown that pseudorotational barrier is \(\approx 4.3 – 8.4\) kJ mol\textsuperscript{-1} for imidazole nucleosides and \(\approx 16.7\) kJ mol\textsuperscript{-1} for their natural analogue. A recent study\textsuperscript{52} showed that the conformational transition pathways between N-type and S-type pseudorotamer of 2'-deoxyadenosine is 9.2 ± 0.8 kJ mol\textsuperscript{-1} using stochastic difference equation (SDE) algorithm. However, most of the theoretical calculations on pentofuranose have been performed assuming a constant puckering amplitude (\(\Psi_m\)), taken as the mean value from crystallographic data. This assumption lowers the reliability of theoretical data, most probably\textsuperscript{53} increasing the height of pseudorotational barrier.

The \(^{13}\text{C}\) relaxation measurement\textsuperscript{54} for purine and pyrimidine ribonucleosides in liquid deuteroammonia between +40° C and -60° C showed that the barrier of pseudorotation for purine nucleosides is 18.7 ± 2.0 kJ mol\textsuperscript{-1} and for pyrimidine nucleosides even greater, \(\approx 25\) kJ mol\textsuperscript{-1}. The temperature-dependent \(^2\text{H}\) and \(^{13}\text{C}\) relaxation study\textsuperscript{55} on selectively deuterated thymidines and allofuranos as well as their comparisons with the conformationally constrained analogues and abasic sugars failed to determine the activation energy barrier of pseudorotation. This is because of the fact that the internal motions are heavily coupled with the overall molecular reorientations, which prevents dissection of the observed activation energy barrier of 20 – 23 (\(\pm 0.9\)) kJ mol\textsuperscript{-1} into contribution from pseudorotational interconversions, rotation around the glycosyl and C4'-C5' torsions and overall tumbling. Nevertheless, this experimental estimate\textsuperscript{55}
gives an upper limit for the pseudorotational barrier. However, studies on the solid-state deuterium NMR line shape analysis\textsuperscript{56} of [2\textsuperscript{"-}D\textsubscript{2}]-2\textsuperscript{\prime}-deoxycytidine at the position C3 in selectively deuterated DNA [d(CGCGAATTCGCG)]\textsubscript{2}, using a double-well potential, have shown the magnitude for pseudorotational barrier of furanose as 13.8 kJ mol\textsuperscript{-1}.

2.2 Types of stereoelectronic forces and their interplay

Through our solution NMR studies, we have attempted to show the dynamic character of the interdependency of the electronic nature of the aglycone and the sugar conformation, dictating the phosphate backbone torsions. The important aspect of the dynamic interdependency of the aglycone-sugar-phosphate orientation in nucleic acids is that it can be modulated by the change of the environment with a certain energy penalty. The mechanism of this modulation is stereoelectronic in character. Through our earlier studies, we have explored the nature of the stereoelectronic forces (Figures 1A and 1B) arising from the anomeric interaction \([nO4' \rightarrow \sigma^*_{C1'-N1/9}]\) orbital mixing, Figure 1A\textsuperscript{36,59,62,63,66,70,75} and gauche interaction \([\sigma_{C-H} \rightarrow \sigma^*_{C-O}]\) Orbital mixing, Figure 1B\textsuperscript{7,60,61,65,68-70,74,75}. It has also been shown that this stereoelectronic effect could be tuned by choice of the sugar substituents\textsuperscript{7,65,75} and their ionization state\textsuperscript{7,68,75} as well as by their complexation with potential ligands present in the medium. This process
specific conformations in a predictable manner in nucleos(t)ides by having appropriate substituent(s) in the sugar moiety. The strength of these stereoelectronic forces depends on the appropriate overlap of donor and acceptor orbitals; the stereoelectronic effects induced stabilization ($\Delta E_s$) is proportional to the square of the overlap ($S$) of the donor and acceptor orbitals, and is inversely proportional to their energy difference ($\Delta E_{orb}$): $\Delta E_s \propto S^2 / \Delta E_{orb}$.

The aglycone promoted AE (Figure 1A) drives the sugar conformation towards the N-type pseudorotamers. The 3'-substituent [OX, where X = H (Figure 1B) for natural $\beta$-$D$-nucleosides and X = PO$_3$H$^-$ for $\beta$-$D$-nucleotides] promoted GE[O4'-C4'-C3'-O3'] leading to $\sigma_{C3'-H3'} \rightarrow \sigma^*_{C4'-O4'}$ orbital mixing

[6,65,68,70,74,75] steers the sugar conformation towards the S-type geometry.65,74 However, 2'-OH in natural RNA has been involved in gauche interaction66 with both O4' of the pentofuranose [i.e. O2'-C2'-C1'-O4' leading to $\sigma_{C2'-H2'} \rightarrow \sigma^*_{C1'-O4'}$ orbital mixing6,68,70,74,75] and N9/1 of aglycone [i.e. O2'-C2'-C1'-N9/1 leading to $\sigma_{C2'-H2'} \rightarrow \sigma^*_{C1'-N9/1}$ orbital mixing70,75]. It has been found that GE[O2'-C2'-C1'-O4'] and GE[O2'-C2'-C1'-N9/1] are counteractive; the former drives the pseudorotational equilibrium of the sugar moiety toward N-type conformation, whereas the latter steers it towards S-type conformation.

2.2.1 Tunable anomeric effect
The overall effect of nucleobase is dictated by its electronic make-up

[7,68,70] as well as the sugar substituents7,66,75. The pairwise comparison (see Tables 6 and 7 in ref 7) of 2',3'-dideoxynucleosides ($\beta$-$D$-ddN), 2'-deoxynucleosides ($\beta$-$D$-dN) and ribonucleosides ($\beta$-$D$-rN) shows that pyrimidine aglycone has larger AE than purine counterpart [i.e. $G < A < T \approx U < C$] at N-state.

The electronic make-up of nucleobase changes as a result of protonation (P) and/or deprotonation (D) of nucleobase,68,70 [compared to the neutral (N) counterpart] which results in changes of the strength of the AE. Lower electron density at N9/1 of the protonated nucleobase enhances nO4' $\rightarrow \sigma^*_{C1'-N1/9}$ Orbital interaction (strengthening AE increase the preference for N-type sugar7,68,70), whereas the reverse is true for deprotonation. Thus the pH-dependent (depending upon P- or D-state compared to the N-state) tunibility of the N $\leftrightarrow$ S equilibrium: $\Delta \Delta G^o = [\Delta G^o]_P$ or D $- [\Delta G^o]_N$ for $\beta$-$D$-rN are as follows (see Table 2 of ref. 7): [Guanosine]$_{P-N}$ (3.0 kJ mol$^{-1}$ for N-type) $> [Guanosine]_{D-N}$ (1.3 kJ mol$^{-1}$ for S-type) $> [Adenosine]_{P-N}$ (1.3 kJ mol$^{-1}$ for N-type) $> [Cytidine]_{P-N}$ (0.2 kJ mol$^{-1}$ for N-type) $\approx [Uridine]_{D-N}$ (0.4 kJ mol$^{-1}$ for S-type).

The strengthening of AE in 8-Aza-3-deazaguanine76 and 7-deaza-2'-deoxyxynucleosides77 (stabilizing N-type conformation) compared to guanine counterpart due to the redistribution of electron-density from N9 into the fused pyrimidine moiety. The pH dependent $^1$H NMR studies with carbocyclic nucleoside73 (such as aristeromycin where O4' is replaced with
CH₂), established the role of AE in the drive of the sugar conformation in N- and C-nucleosides. The S-C-N promoted AE in 4'-thiodeoxynucleosides occurs in the following order: thymine < cytosine < guanine < adenine. This trend is the opposite compared to those of natural nucleosides. Moreover, S-C-N promoted AE in 4'-thiodeoxynucleosides is weaker than the O-C-N promoted AE in the corresponding natural counterpart (4'-oxonucleosides).

The pH-dependent conformational analyses by ¹H NMR of C-nucleosides showed that the participation of nO₄ → σ⁺C₁'-C₅(sp²) orbital mixing (AE) in C-aglycone driving its N ⇌ S equilibrium. The comparison of C- and N-nucleosides showed that the anomic tunability is more pronounced in pyrimidine C-nucleosides than that of pyrimidine N-nucleosides, whereas there is no significant difference between purine C- and N-nucleosides.

2.2.2 Tunable gauche effects
Conformational analyses with of 2'-deoxy-2'-substituted uridine and adenosine derivatives showed qualitatively that the population of N-type pseudorotamers linearly increases with increasing electronegativity of the 2'-substituent as a result of the enhanced 2'-GE[O₂'-C₂'-C₁'-O₄']. Similarly, the two-state N ⇌ S equilibrium in 2'-methylthionucleosides is strongly biased toward S-type conformations in CD₃OD, and the effect of 2'-SMe has been attributed both to its reduced electronegativity (i.e. resulting in weaker GE[S₂'-C₂'-C₁'-O₄'] and GE[S₂'-C₂'-C₁'-N1/9]) and increased steric bulk (resulting in the destabilization of N-type pseudorotamers).

Similarly, electronegative nature of 3'-substituent (X) dictates 3'-GE[X₃'-C₃'-C₄'-O₄']. The temperature-dependent pseudorotational analyses of 3'-substituted (X)-β-D-ddN [where X = H, NH₂, OH, OMe, NO₂, OPO₃H⁻ and F] showed the increasing preference for gauche orientation within [X₃'-C₃'-C₄'-O₄'] fragment with increasing electronegativity 3'-substituents [H < NH₂ < OH < OMe < NO₂, OPO₃H⁻ < F].

2.2.3 Configuration dependent gauche effects
The GE of highly electronegative fluorine (F) substituent, has a profound stereoelectronic effect, thereby it governs the overall conformation of the sugar ring. The sugar moieties in 2''-α-fluoro-2',3'-β-
D-dideoxyuridine (2''-α-FddU) and 3'-β-fluoro-2',3'-β-D-dideoxyuridine (3'-β-FddU) adopt exclusively N-type conformations\(^8\) (Figure 2), owing to the cooperative drive of the 2''(α)F-GE in 2''-α-FddU and 3'((β)F-GE in 3'-β-FddU, respectively with the AE. In contrast, as a result of the configuration-dependent GE (Figure 2), the two-state N ⇔ S pseudorotational equilibrium in 2'-β-fluoro-2',3'-β-D-dideoxyuridine (2'-β-FddU) and 3''-α-fluoro-2',3'-β-D-dideoxyuridine (3''-α-FddU) are strongly biased to the S-type conformers because of the predominance of the 2'(β)F-GE in 2'-β-FddU and 3''(α)F-GE in 3''-α-FddU respectively (Figure 2), over the AE.\(^8\)

2.2.4 The interplay of competing anomeric and gauche effects

The preferred sugar vis-à-vis backbone conformation of nucleic acids depends on the culmination of the mutual interplay of substituent (i.e. electronic nature of the aglycone\(^7,68,70\) as well as those of other substituents on the furanose ring\(^7,66,75\)) dependent AE and GE. The interplay between AE and GE in β-D-dN has been schematically shown in Figure 3. The participation of stronger 3'-GE weakens AE in β-D-dN compared to β-D-ddN and β-D-rN. The weaker AE in β-D-dNs compared to β-D-ddNs (in N-state) is explained\(^7\) as follows: the electron-withdrawing character of 3'-OH reduces the electron density around O4' in β-D-dNs compared to β-D-ddNs, making O4' lonepair in the former less available for n\(_{O4'}\) → σ\(_{C1'-N9}\) interactions compared to the latter (Figure 3). Thus, the electron density around O4' is maximal in abasic sugar and as a result σ\(_{C3'-H3'}\) → σ\(_{C4'-O4'}\) interaction is disfavored (because of higher difference between their energy levels and therefore relatively poorer overlap and orbital mixing) compared to that in β-D-dNs. However, the effect of a nucleobase in β-D-rNs is the same as in β-D-ddNs at the neutral
state\textsuperscript{7}, owing to the partial cancellation of the GE[O2'-C2'-C1'-O4'] (steer towards N-type) and GE[O3'-C3'-C4'-O4'] (steer towards S-type) in the former. Additionally, the GE[O2'-C2'-C1'-N1/9] involving $\sigma_{C2'-H2'} \rightarrow \sigma^*_{C1'-N9/1}$ orbital mixing drives the pseudorotational equilibrium of the sugar moiety in $\beta$-D-ribonucleos(t)ides toward S-type pseudorotamers. The strength of GE[O2'-C2'-C1'-N1/9] is in the order of Adenine (-7.9 kJ mol\textsuperscript{-1}) > Guanine (-6.7 kJ mol\textsuperscript{-1}) > Cytidine(-2.5 kJ mol\textsuperscript{-1}).\textsuperscript{7,70}

Moreover, the ability of 2'-OH to counteract the GE[O3'-C3'-C4'-O4'] by stabilizing N-type conformations increases in the order\textsuperscript{7}: 3'-deoxyadenosine > adenosine > adenosine-3'-monophosphate which can be attributed to the more efficient GE[O2'-C2'-C1'-O4'], or to the weakening of the GE[O2'-C2'-C1'-N9]. However, the GE[O4'-C4'-C3'-O3'(PO\textsubscript{2}H\textsuperscript{-})] is found to be less efficient\textsuperscript{7} by -1.4 kJ mol\textsuperscript{-1} in adenosine-3'-ethylphosphate than 2'-deoxyadenosine-3'-ethylphosphate, thereby showing the stereoelectronic modulation of 2'-OH in the former besides its intra- and intermolecular hydrogen bonding capability (Paper IV).

So far the studies on the stereoelectronic effects in nucleosides and nucleotides showed that 2' (or 2") and/or 3' (or 3") F/O mediated GE is stronger than AE. A recent report\textsuperscript{86} from Seela's group showed, for the first time, that the AE can be stronger than the F/O mediated GE in dictating the furanose conformation of the nucleosides. Thus, they have shown\textsuperscript{86} that they could overcome the strong GE of 2'-F substituent in the $\beta$-face (ara configuration) by the introduction of the electron withdrawing Br atom in the pyrazolo [3,4-d]pyramidine aglycone. The two nucleosides, 3-bromopyrazolo [3,4-d]pyrimidine-2'-deoxy-2"-flouro and its ara counterpart showed exclusive N-type conformation in solution and solid state,\textsuperscript{86} which is unexpected in view of the earlier studies with 2'($\beta$)-flouro substituted nucleosides (e.g. 2'- flies seen in section 2.2.3). The strong electron withdrawing nature of nucleobase in these nucleosides enhances the $n_{O4*} \rightarrow \sigma^*_{C1'-N1}$ interaction (AE) which prevails over the gauche interactions, thereby dictating the total sugar pseudorotational drive toward the C3'-endo conformation in solution.

### 2.3 Biological implications of stereoelectronic effect in nucleic acids

The importance of the recognition of pentofuranosyl-sugar conformation by various enzymes has been demonstrated in various studies. Methanocarba nucleosides\textsuperscript{87,88} having a rigid bicyclo[3.1.0]-hexane template have been instrumental in defining the role of sugar pucker by stabilizing biologically preferred sugar conformation [e.g. (N)-(−)-methanocarba-A is preferentially recognized as substrate of adenosine deaminase\textsuperscript{87}]. It has also been shown\textsuperscript{88} that (N)-methanocarba-ATP with N-type sugar conformation was 138- and 41-fold more potent at recombinant human P2Y\textsubscript{1} and P2Y\textsubscript{2} receptors as antagonists, than racemic (S)-methano-ATP with S-type sugar conformation.
An NMR study with $^{13}$C/$^2$H double-labelled 2'-deoxyadenosine (dAdo) and the corresponding 2'-deoxycytidine (dCyd) moieties in the complexes with human recombinant deoxycytidine kinase (dCK) showed that the ligands (i.e. dCyd and dAdo) adopts a S-type sugar conformation when bound to dCK endowing the importance of sugar conformation in forming kinetically favoured enzyme complex.

Recent work from our laboratory on RNase H cleavage of the antisense oligonucleotide (AON)/RNA hybrid duplex incorporated with conformationally constrained (in N-type or N/E-type) nucleoside [1-(1',3'-O-anhydro-β-D-psico-furanosyl)thymine/cytosine], showed that local conformational changes transmit up to the total of 5 neighboring nucleotide residues including the modification which has been recognized by the cleavage. The incorporation of 2'-deoxy-2'-fluoro-β-D-nucleoside having preferred O4'-endo sugar conformation in a DNA/RNA hybrids duplex invokes improved nuclease resistance and RNase H digestion of DNA/RNA hybrid duplex, indicating the importance of 2'-GE directed preferred sugar N/E-type conformation in duplex stability as well as RNase H recognition.

3. Present Work (Papers I – III)

The intrinsic dynamics and architectural flexibility of nucleic acids resulting into specific function are the net outcome of cooperative interplay of pentofuranose, nucleobase and phosphodiester moieties. In oligonucleotides the protonation, deprotonation and/or methylation of nucleobase directly affects their hydrogen-bonding capabilities (and other electrostatic interactions) and therefore induces change in overall three-dimensional structure. Thus, the structural modification of natural nucleotides can influence the biochemical as well as functional aspects of nucleic acids in general.

3.1 Aglycone-sugar-phosphate conformational cooperativity

In this study (Papers I and II) we have shown that in absence of intramolecular base-base stacking, the change of the electronic character of aglycone through protonation (i.e. N7 of guanin-9-yl, N1 of adenin-9-yl and N3 of cytosin-1-yl) not only modulates the shift of N ⇆ S pseudorotational equilibrium of their constituate sugar by strengthening the anomeric effect but is also transmitted to the sugar-phosphate backbone to steer the conformation of backbone. The 3'-ethylphosphate, 5'-methylphosphate derivatives (Scheme 4) of guanosine (MepGpEt, 1a), adenosine (MepApEt, 1b) and cytosine (MepCpEt, 1c) in conjunction with their abasic counterpart [Etp(ab)Me, 1d] can be considered as a mimicking model of trinucleoside
diphosphate [Scheme 1]. These studies have revealed three aspects in the aglycone-sugar-phosphate conformational cooperativity (sections 3.1.1 to 3.1.3):

3.1.1 pD-dependent shift of N ⇔ S pseudorotational equilibrium and ε ⇔ ε' equilibrium in nucleotide.

The N ⇔ S pseudorotational equilibrium in 1a and 1b is gradually shifted towards N-type pseudorotamers [79% S (for 1a) and 76% S (for 1b) in the N-state to 55% S (for 1a) and 67% S (for 1b) in the P-state] as reflected from change of the free energy of the N ⇔ S pseudorotational equilibrium \( \Delta G^\circ (N ⇔ S) \), at 298 K from -3.3 kJ mol\(^{-1}\) (for 1a) and -2.8 kJ mol\(^{-1}\) (for 1b) in the N-state to -0.1 kJ mol\(^{-1}\) (for 1a) and -1.7 kJ mol\(^{-1}\) (for 1b) in the P-state [Figure 1 in Paper I and Figure 2 in Paper II].

The \(^1\)H NMR analysed conformational bias (in terms of two-state \( \varepsilon^+ ⇔ \varepsilon^- \) equilibrium\(^7\)) across C3'-O3' bond (ε torsion) showed that the population of \( \varepsilon^- \) (at 298K) decreases in N-state compared to that in P-state. The corresponding shift of the free energy of \( \varepsilon^+ ⇔ \varepsilon^- \) equilibrium (\( \Delta G^\circ (\varepsilon^+ ⇔ \varepsilon^-) \), at 298 K) is from -2.1 kJ mol\(^{-1}\) (for 1a) and -1.9 kJ mol\(^{-1}\) (for 1b) in the N-state to +0.3 kJ mol\(^{-1}\) (for 1a) and -1.5 kJ mol\(^{-1}\) (for 1b) in the P-state respectively. Interestingly, the sigmoidal plot of pD-dependent \( \Delta G^\circ \) values for the N ⇔ S
as well as $\varepsilon^{\dagger} \Rightarrow \varepsilon^{-}$ equilibrium in 1a and 1b has $pK_a$ almost identical to that of constituent guanin-9-yl and adenin-9-yl nucleobase, as determined independently from the plot of pD-dependent $\delta$H8 and $\delta$H2 [Figure 1 in Paper I and Figure 2 in Paper II].

3.1.2 The cooperative shift of the $(N, \varepsilon^{\dagger}) \Rightarrow (S, \varepsilon^{-})$ equilibrium in nucleotides

The cooperative shift of the $(N, \varepsilon^{\dagger}) \Rightarrow (S, \varepsilon^{-})$ equilibrium as the result of protonation of aglycone in 1a and 1b is evidenced by the straight line obtained from the plot of pD-dependent $\Delta G^\circ_{(N=S)}$ as a function of pD-dependent $\Delta G^\circ_{(\varepsilon^{\dagger}=\varepsilon^{-})}$ [Figure 4] as well as from plots of $\Delta G^\circ_{(N=S)}$ and $\Delta G^\circ_{(\varepsilon^{\dagger}=\varepsilon^{-})}$ as a function of both $\delta^{31}$P and $\delta^{1}$H [Figures 2 and 5 in Paper I and Figures 3 and 4 in Paper II].

As a control experiment, the difference in $^3J_{HH}$, $^3J_{HP}$ and $^3J_{CP}$ coupling constant values between neutral and acidic pDs at 298K was found to be negligible in abasic phosphodiester 1d, showing that owing to the absence of aglycone, the $N \Rightarrow S$ equilibrium as well as conformation across C3'-O3' remain unbiased at all pDs compared to 1a and 1b.

3.1.3 Tunability of aglycones and transmission of the electronic character

These studies therefore show a complete interdependency of conformational preference of sugar and phosphate backbone (in absence of intramolecular base-base stacking) as the protonation $\Rightarrow$ deprotonation equilibrium of the aglycone changes as a function of pH. However, this aglycone dependent conformational transmission till sugar-phosphate backbone via pentofuranose depends upon the tunability of aglycone vis-à-vis conformational modulation of sugar geometry. Our control studies with 1c at the N- and the P-states showed that the relative conformational tunability [Figure 5] is in order: MepGpEt (1a) > MepApEt (1b) > MepCpEt (1c). Moreover, this tunable transmission, when compared to the abasic
counterpart, is found to be much stronger at the 3'-phosphate compared to the 5'-end (see Figure 5 in Paper I and Figure 4 in Paper II).

Figure 5. The relative conformational tunibility of MepGpEt (1a), MepApEt (1b) and MepCpEt (1c) between neutral (N) and protonated (P) state at 298 K. Panel (A) shows the relative change of chemical shift of aromatic protons $[\Delta \delta_{(P - N)}]$ in 1a – c; Panel (B) and (C) show the relative change of free energy of the N ⇆ S pseudorotational equilibrium ($[\Delta \Delta G^\circ_{(N \leftrightarrow S)}]_{P - N}$, in kJ mol$^{-1}$) and that of $\varepsilon^+$ ⇆ $\varepsilon^-$ equilibrium ($[\Delta \Delta G^\circ_{(\varepsilon^+ \leftrightarrow \varepsilon^-)}]_{P - N}$, in kJ mol$^{-1}$) in 1a – c. The relative order of tunibility is 1a > 1b > 1c.

3.1.4 Mechanistic overview of RNA as molecular wire for conformational transmission (Papers I)

In earlier works with 3a – 3e and 3f – 3j (Scheme 4) at the neutral pH, the methylene protons of the 3’-ethyphosphate moiety of ribo analogues (3a –3e) have been found to be non-equivalent, which slowly become isochronous (similar to 2’-deoxy counterparts) at high temperature ($\geq$ 348 K). This is attributed to the absence of 2’-OH promoted intramolecular 2’OH … O3'-P H-bonding (see Paper IV for further details).

Our studies with 1a and 1b showed similar temperature-dependent multiplicities of the methylene protons of 3'-ethylphosphate moiety at 1.0 $\leq$ pD $\leq$ 6.7, thereby suggesting that the 2’OH···O3'-P H-bonding remains the same in 1a and 1b over the whole pD range at room temperature (298 K). Thus, all changes of the pD-dependent free energies observed at 298 K for 1a and 1b can be attributed to the changes of the protonation ⇆ deprotonation equilibrium of the aglycones to drive the sugar-phosphate backbone in a concerted manner.

Figure 3 and ref. 12 in Paper I have presented a detailed discussion of the molecular orbital diagram (taking 1a as a model) based interpretation of this concerted conformational transmission. As the pD-tunable change of the electronic character of the nucleobase tunes the strength of the AE, an increased preference of the N-type sugar conformation is imposed because of enhanced $n_{O4'}$ → $\sigma^*_{C1'-N9}$ orbital interaction, which, in turn, affects the strength of the 3'-GE [O3'-C3'-C4'-O4'] by retuning the energy levels of the donor and the acceptor orbital in the $\sigma_{C3'-H3'}$ → $\sigma^*_{C4'-O4'}$ interaction. The extent
Figure 6. The relative donor and acceptor abilities of various orbitals in the N- and the P states are shown. Since the electronic state of the aglycone modulates the sugar conformation which in tum modulates the phosphate torsion, the $1\text{n}_{sp^2}$ (p-type, O3') orbital is placed at a relatively lower energy level than $1\text{n}_{sp^2}$ (p-type, O4'). As the AE involving O4'-C1'-N9 starts operating, the strength of the 3'-GE [O3'-C3'-C4'-O4'] (σC3'-H3' → σ*C4'-O4') counteracts the AE owing to the relatively lower energy of σ*C4'-O4' [A_N becomes A_P state with the protonation of the nucleobase]. As the nucleobase becomes protonated, σ*C1'-N9 becomes a better acceptor and the O4'-C1'-N9 AE is strengthened, and that makes 3'-GE more effective [σC3'-H3' (BN becomes BP state) → σ*C4'-O4', i.e. more effective orbital mixing of BP with AP, see $\Delta\Delta H^\circ_{10}$ in Table 6 of ref. 7]. However, the AE is stronger than the 3'-GE, therefore we see overall stabilization of more N-type sugars in the P-state ($\Delta\Delta E(\text{GE}) < \Delta\Delta E(\text{AE})$, ref. 7] compared to that in N-state. The AE involving O3'-P3'-O(ester) is weaker in the S-type pseudorotamers (at the N-state, i.e. in C_N state) than in the N-type counterparts (at the P-state, i.e. in C_P state) because the σC3'-H3' orbital overlaps with the σ*C4'-O4' (i.e. 3'-GE is stronger in the former state), reducing the electron density at O3' (B). This means that the $1\text{n}_{sp^2}$ (p-type, O3) is relatively less available in S-type conformation (at the neutral pH) to interact with the σ*C3'-O(ester) in S-type conformation of the nucleobase, which shows that the O3'-P3'-O(ester) AE is weaker in S-type conformation.
of $\sigma_{C3'-H3'} \rightarrow \sigma^*_{C4'-O4'}$ participation influences the electron density at O3’, which in turn modulates the AE involving O3’-P3’-O(ester) in a concerted manner. Figure 6 shows that the corresponding energy levels of the orbitals involved in the AE and GE, on basis of their relative acceptor/donor abilities in a purely qualitative manner. This translates itself in terms of relative strength of GE and AE and the preferred conformational states, which make the RNA to act as a molecular wire. The flow of electronic modulation from the nucleobase till the phosphate via sugar moiety is reflected by the fact that the hybrid orbital produced by the AE involving O3’-P3’-O(ester) is at a lower energy level than the corresponding hybrid orbital resulting from AE involving O4’-C1’-N9. However, the final proof of the operation of AE [O3’-P3’-O(ester)] could only be experimentally obtained if we could only measure the $\zeta$ and $\alpha$ torsions across the 3’-phosphate backbone and the preferential O3’-P3’-O bond angle.

3.2 Strength of 3’-Gauche effect dictates the EF-Tu recognition of aminoacyl-tRNA: studies based on the mimicking model (Paper III)

The aminoacylation at 3’-terminal (acceptor arm, Scheme 5) of tRNA and its subsequent recognition by the Elongation Factor Tu (EF-Tu) is an important step during the process of in vivo protein synthesis.99,100 The EF-Tu is a guanine nucleotide binding protein factor that, when complexed (EF-Tu*GTP) with guanosine 5’-triphosphate (GTP), binds elongator aminoacyl-tRNAs (aa-tRNAs).99-101 Tight binding of a tRNA by EF-Tu requires the presence of cognate amino acid esterified to its adenosine of 3’-terminal CC$\Delta_{\text{OH}}$ by appropriate aa-tRNA synthetase.102 The discrimination between correctly and incorrectly charged aa-tRNA by EF-Tu binding clearly shows99,101 its specificity in such recognition process.

Anthranilic acid charged yeast tRNA$^{\text{Phe}}$ or E. coli tRNA$^{\text{Val}}$ form a stable complex with EF-Tu*GTP, hence the 2’- and 3’-O-anthraniloyladenosines and their 5’-phosphate counterparts (2a – d, Scheme 4) have been conceived to be the smallest units that are capable to mimic aa-tRNA.103-106 Since 2c and 2d binds more efficiently105 to EF-Tu*GTP complex compared to 2a and 2b respectively, we delineated the stereoelectronic features that dictate the conformation of former vis-à-vis lattar as well as addressed how their
structures and thermodynamic stabilizations are different from Ado (4a) and 5'-AMP (7b).

2c ($\Delta G^\circ_{(N=S)} = -4.6 \text{ kJ mol}^{-1}$) and 2d ($\Delta G^\circ_{(N=S)} = -3.9 \text{ kJ mol}^{-1}$) have relatively more stabilized S-type conformation, whereas the $\Delta G^\circ_{(N=S)}$ for 2a and 2b are -0.9 and -1.8 kJ mol$^{-1}$ respectively, suggesting that the 3'-GE [H3'-C3'-C4'-O4'] of 3'-O-anthraniloyl group is stronger than 2'-GE [H2'-C2'-C1'-N9'] of 2'-O-anthraniloyl in the drive of the sugar conformation to S-type. Since the EF-Tu can specifically recognize the aminoacylated-tRNA from the non-charged tRNA, we have assessed the free-energy ($\Delta G^\circ$) for this recognition switch ($[\Delta G^\circ_{(N=S)}]_{2a} - [\Delta G^\circ_{(N=S)}]_{7b}$) to be at least $\approx -2.9 \text{ kJ mol}^{-1}$. This specific recognition process requires 3'-terminal adenosine (after aminoacylation) to move to the hydrophobic pocket.$^{106,107}$ This would mechanistically require destacking from acceptor helix, likely assisted by S-type sugar pucker of 2c or 2d. Thus, the antibiotic puromycin$^{108}$ with N-type sugar moiety, being an excellent aa-tRNA-mimic and a powerful inhibitor, it although failed to interact with EF-Tu.

The 3'-O-anthraniloyl compounds 2c and 2d are more flexible than the isomeric 2'-O- counterparts 2a and 2b as evident from the temperature-dependent $^3J_{H,H}$ analysis (Table 2 in Paper III). The thermodynamics of the transacylation ($\Delta G^\circ < 0$ for 2'→3' and $\Delta G^\circ > 0$ for 3'→2') reaction of 2a $\rightleftharpoons$ 2c ($\Delta G^\circ = -1.2 \text{ kJ mol}^{-1}$) and 2b $\rightleftharpoons$ 2d ($\Delta G^\circ = -1.7 \text{ kJ mol}^{-1}$) is cooperatively dictated by the N $\rightleftharpoons$ S pseudorotational equilibrium of their sugar moiety, which in turn is controlled by a balance of the 3'- vis-à-vis 2'-GE. This also explains the slower 3'→2' transacylation rate ($1 - 4 \text{ s}^{-1}$) compared to 2'→3' transacylation ($3 - 11 \text{ s}^{-1}$) as found in earlier kinetic studies.$^{109}$

3.3 Implications

The thermodynamics of conformational transmission in mononucleotides clearly show that the conformational transmission across the nucleotidyl wire (i.e. from aglycone to sugar to phosphate) is responsible for modulation of the sugar-phosphate backbone as a result of change of aromatic character of the aglycone. The physico-chemical roles of the aglycone, sugar and phosphate, depending upon the local microenvironment, dictate the stereoelectronic forces thereby affecting the function as well as the self-assembly of nucleic acids. It is thus clear that any intermolecular interaction between nucleic acid and a ligand is expected to produce similar effects as arising from the protonation or deprotonation of the aglycone.

Small organic molecules can be designed as tRNA-mimics$^{110}$ (like puromycin$^{108}$) to manipulate the functional properties of tRNA. Thus, 3'-O-anthraniloyl adenosine derivatives can be used for the EF-Tu recognition mimicry with the help of this stereoelectronic tuning to induce the recognition switch.
4. The contribution of H-bonding in biomolecular interactions

4.1 Nature of hydrogen bonding

Conventional H-bond interaction (X-H…Y, Scheme 6) involves two electron-withdrawing atoms (X and Y), one being attached to hydrogen (X, the donor) and other bearing lone electron pair (Y, the acceptor).\textsuperscript{111-113} The H-bonding is a fundamental feature of chemical structure and reactivity. Although the precise definition and nature (whether electrostatic, charge transfer or dispersion) of H-bonding continues to be elusive,\textsuperscript{114} it can be roughly classified,\textsuperscript{112,115,116} depending on their strength (in terms of the enthalpy of H-bonding): (i) strong [\textit{e.g.} F-H…F in gas phase; 24 – 40 kcal mol\textsuperscript{-1} for single-well H-bond (SWHB) and 12 – 24 kcal mol\textsuperscript{-1} for low barrier H-bond (LBHB)] (ii) moderate [\textit{e.g.} O-H…O in water, alcohol and monoanion of dicarboxylic acid like hydrogen phthalate; 4 – 12 kcal mol\textsuperscript{-1}] and (iii) weak [\textit{e.g.} C-H…O,\textsuperscript{117,118} N-H…\pi\textsuperscript{119} and non-linear H-bonds\textsuperscript{120,121}; < 4 kcal mol\textsuperscript{-1}].

The strength of H-bond correlates strongly with H-bonding length (r\textsubscript{HY}, Scheme 6),\textsuperscript{122} however, little or no correlation has been found both experimentally\textsuperscript{122} and theoretically\textsuperscript{124} with H-bond angle (\angle XHY, Scheme 6) over the range of 180 ± 30°. Angle bending beyond ± 30° can lead to weakening of H-bonding.\textsuperscript{125} Stronger H-bonds are those where the donor and acceptor has similar pK\textsubscript{a} values ("pK\textsubscript{a} match", \textit{i.e.} \Delta pK\textsubscript{a} = 0, Scheme 6), and that allows the donor and acceptor to share the proton equally which has been evidenced from both experimental as well as theoretical studies in the literature.\textsuperscript{126-129} A linear correlation has been found\textsuperscript{128} between H-bond strength (log K\textsuperscript{HB}) and \Delta pK\textsubscript{a} for homologous series of substituted salicylic acids in both DMSO and aqueous solvent.

In aqueous environment, due to competition with solvent, strong H-bonding is relatively less abundant. Thus, weak and moderately strong H-bonding (commonly found O-H…O,\textsuperscript{130,131} O-H…N,\textsuperscript{132,133} N-H…O\textsuperscript{134,135} and N-H…N\textsuperscript{112,136}) contributes significantly to the structure, properties and recognition pattern of biomolecules like carbohydrate, protein and nucleic acids. The directive power of H-bonds is apparently one of the major factors for the self-assembly and specificity of biopolymer structures. Jeffrey & Saenger have concluded\textsuperscript{112} that the energy of conventional H-bonds in biological system ranges between 1 – 4 kcal mol\textsuperscript{-1} depending upon the electronic character of donors and acceptors. However, recent studies from
Frey et al. showed\textsuperscript{116} the importance of LBHB in transition state stabilization of the enzymatic complex. Studies\textsuperscript{137} with Serine Protease inhibitors have recently elucidated the presence of multi-centered short H-bond arrays.

4.1.1 \textit{Hydrogen bonding in nucleic acids}

The H-bonded base pairing is one of the forces for duplex stabilization. The strength of base pairing is found\textsuperscript{1,2} to be -0.5 to 1.5 kcal mol\textsuperscript{-1} based on the sequence dependent competition between H-bonding and stacking. X-ray and NMR studies showed Watson-Crick basepairing\textsuperscript{1,2} in usual A- and B-type RNA and DNA.\textsuperscript{1,2,5} However, non-Watson-Crick basepairing have also been found\textsuperscript{1,2,138,139} in several nucleic acid structures. Wobble basepairing theory\textsuperscript{140} has been proposed to explain the degeneracy of triplet codon. Three base coplanar interactions via H-bonding have been found in triple-stranded DNA\textsuperscript{141}. Studies by Seela and colleagues have shown the importance of reverse Watson-Crick basepairing\textsuperscript{142} by isoguanosine and isocytidine to form the stable parallel stranded DNA duplex. Recent studies\textsuperscript{143,144} have demonstrated the role of tertiary H-bonding in stabilization of RNA structures\textsuperscript{143} as well as backbone mediated interresidue N-H···O H-bonding in PNA:RNA heterduplex.\textsuperscript{144} Besides, H-bonding involving 2'-OH has also been elucidated (sections 4.2 and 5.1).

4.1.2 \textit{Hydrogen bonding in peptides and carbohydrates}

Major contribution to catalysis by many enzymes are provided by formation of the H-bonds from catalytic core of enzyme to the bound substrate like in mandelate racemase,\textsuperscript{132,134} triosephosphate isomerase,\textsuperscript{145} citrate synthetase,\textsuperscript{146} ketosteroid isomerase,\textsuperscript{131,146} chymotrypsin and other serine protease\textsuperscript{134,147} as well as in myoglobin ligation\textsuperscript{148} with heme protein. In proteins the energetic contributions of individual hydrogen bonds are often assessed through deletion studies. These estimates\textsuperscript{149-154} are generally in the range ~2 – 5 kcal mol\textsuperscript{-1}. Several studies, based on variable-temperature NMR and IR in aprotic solvents to measure the contribution of H-bonding in protein folding, have been performed using model compounds: homologous diamiades,\textsuperscript{155} \(\beta\)-alanine derivative of phenoixanthin derivatives\textsuperscript{156} and alkyl substituted \(\beta\)-amino acid containing polyamides.\textsuperscript{157} The NMR and molecular modeling studies using hydroxyl proton in conformational analysis of carbohydrates\textsuperscript{121,158-161} have also reported the presence of hydrogen bonded interactions.

4.2 \textit{Hydrogen bonding by 2'-hydroxyl group in RNA}

The 2'-OH distinguishes\textsuperscript{1,162,193} RNA from DNA both functionally\textsuperscript{163,165,169,177,183} as well as structurally.\textsuperscript{166,167,174} The 2'-OH group in RNA is a powerful handle to drive the sugar-phosphate backbone conformation both stereoelectronically\textsuperscript{7,67} and by direct interaction\textsuperscript{167} with the neighboring
function, or through intra- and intermolecular hydrogen bonding as well as by inducing differential gradient of hydration.

4.2.1 Importance of 2'-hydroxyl group in RNA
The 2'-OH group is involved in recognition, processing and catalytic properties of RNA, such as the stereospecific transesterification reactions involved in the Group I and Group II splicing reactions,\(^{170,179}\) self-cleavage in lariat-RNA,\(^{180}\) in ribonuclease\(^{181,182,282}\) action and RNA catalysis in ribozyme\(^{165,166,168,173,195,197,300}\) as well as in tRNA processing by RNase P RNA.\(^{184-186}\)

Recent studies\(^{187}\) showed the evidences of the participation of the five 2'-OH groups in tRNA\(^{\text{hhe}}\) to stabilize its complex with EF-Tu from *Thermus thermophilus*. The role of 2'-OH has also been demonstrated\(^{188}\) in the interaction between acceptor stem of *E. coli* tRNA\(^{\text{cys}}\) and cysteine-tRNA synthetase providing high aminoacyl specificity. The substitutions of 2'-OH by either 2'-F, 2'-H, 2'-NH\(_2\) or 2'-OMe group at either U\(^5\), U\(^6\) or C\(^7\) in U\(^5\)U\(^6\)C\(^7\)G tetraloops\(^{189}\) showed the relative change in thermodynamics, which reflects a complex interplay of H-bonding, solvation effect and intrinsic sugar pucker preference. The exchange properties of 2'-OH of a guanosine residue involved in a novel H-bond has been shown\(^{190}\) to contribute to the immobilization of bound AMP by the RNA aptamer. Crystallographic and UV studies showed that incorporation of both ribocytidine (rC) and arabino cytidine (araC) in hexameric d(CGCGCG)\(^{191}\) allows 2'-OH to form intramolecular H-bond with N2 of 5'-guanine by replacing the water bridge in deep groove to stabilize the guanine in syn conformation thereby facilitating the B \(\rightarrow\) Z transition. Recently, thermodynamic studies\(^{192}\) with d(CGCGCG) by incorporating 8-methylguanosine (m\(^8\)rG) compared to 8-methyl-2'-deoxyguanosine (m\(^8\)dG) showed that m\(^8\)rG (with N-type sugar) stabilizes the B \(\rightarrow\) Z transition (even at low salt concentration) more compared to m\(^8\)dG by reducing entropy which arises from hydrophilic 2'-OH in solvent exposed region.

4.2.2 Hydrogen bonding and hydration of 2'-hydroxyl group in RNA
The MD simulation of the tRNA\(^{\text{Asp}}\) anticodon hairpin showed\(^{193}\) that, in C3'-endo sugar pucker, the 2'-OH group can access any of the three orientations: towards (i) the O3', (ii) O4' of the same sugar and (iii) the nucleobase. However, in C2'-endo sugar pucker, the 2'-OH is preferentially directed towards vicinal O3'. Recent studies have found\(^{170}\) that the internucleotidyl 2'-OH of U moiety at the cleavage site (U\(_{i}\)) is indeed intramolecularly H-bonded (2'-OH-H-\(\cdots\)O3') in splicing reaction of the Tetrahymena group I ribozyme. The water bridge model of intermolecular H-bond [2'-O-H-water-\(\cdots\)O3'] as well as intrasrand O2'-H-O4' H-bonding have also been suggested\(^{195}\) by crystal structure analyses of RNA duplex [r(CCCCAGGGG)]\(_2\) at 1.46 Å resolution. The differential hydration pattern of DNA and RNA
duplex showed$^{196}$ relatively more hydrophilic character in minor groove in
the latter (wide, $\sim$11.0 Å)$^{1,2}$ compared to that in the former (narrow, $\sim$5.7
Å).$^{1,2}$ Moreover, the crystal data analyses for DNA-RNA duplex showed$^{197}$
that 2'-OH of a guanosine (rG11 with C3'-endo sugar pucker) residue forms
hydrogen bond to one of the phosphate oxygens from adjacent residue (rC
12). Involvement of 2'-OH mediated H-bonding in the formation of the
ribose zipper motif$^{198,199,200}$ in P4-P6 domain of group I ribozyme, in
A(GC) base triplet formation$^{198}$ and in certain helical stacking in double-
stranded RNA$^{201-204}$ has also been elucidated. Similarly, crystallographic
study$^{205}$ with uridine 3'-monophosphate monohydrate showed an
intramolecular O2'-H … O4' H-bonding in C2'-endo sugar pucker mode. The
crystallographic studies$^{206}$ proposed an intramolecular 2'-O-H … O3' $\leftrightarrow$
3'-O-H … O2' H-bonding in ribonucleosides. However, earlier $^1$H NMR
investigation of 3',5' cAMP in aqueous and mixed solvents showed$^{207}$ the
formation of water bridge by the 2'-OH with vicinal 3'-phosphoryl oxygen.

5. Present Work (Paper IV)

The role of 2'-OH in nucleic acids as proton donor in both intra- and
intermolecular H-bonding in RNA is qualitatively evident from studies
described in section 4.2.2. However, the contradicting results about the
nature of such interaction and absence of any quantitative NMR studies
prompted us to undertake the more detailed studies on the 2'-OH mediated
H-bonding and hydration pattern in RNA at nucleotide level.

5.1 Hydrogen bonding and hydration of 2'-hydroxyl group in nucleosides
and nucleotides

5.1.1 Geometry of intramolecular hydrogen bonding

The NMR constrained molecular modelling (with MM as well as ab initio
methods both in the gas and solution phase) has been used to characterize the
energy minima (Figure 3 and Table 2 in Paper IV) among the four
alternative dihedrals possible from the solution of the Karplus equation for
$^3J_{H2',OH}$ and $^3J_{H3',OH}$ to delineate the preferred orientation of 2'-O-H proton
(Figure 2 in Paper IV) in 3b and 4b as well as for 2'/3'-O-H protons in 4a.
The $\Phi_{H2'-C2'-O2'-H}$ (from NMR constrained ab initio geometry optimization) for
pseudoequatorial 2'-OH with S-type sugar geometry of 3b ($\Phi_{H2'-C2'-O2'-H} =
123.0^\circ$) and 4a ($\Phi_{H2'-C2'-O2'-H} = 133.2^\circ$) in positive transoid domain
corresponds to the closer proximity to the neighboring O3' (Figure 2 in
Paper IV). The presence of intramolecular 2'-OH … O3' H-bonding in 3b and
4a is also corroborated by (i) temperature dependent change in multiplicity
of methylene (-CH$_2$-) protons of ethylphosphate moiety in 3b (see section
A(i) in Paper IV), (ii) the existence of weak long range $^4J_{H2',OH3'}$ in 4a (i.e.
W conformation of H2'-C2'-C3'-O3'-H), (iii) from the preferential orientation of the 2'-OH in 3b and both 2'- and 3'-OH groups in 4a and (iv) solvent polarity studies (see section A(ii) in Paper IV) for 4a. Thus, it has been found that geometrical factors (like bond angle, bending etc.) other than proximity also contribute to the strength of H-bonding.

An alternate possibility of intramolecular H-bonding between 2'-OH and the vicinal non-bridging phosphoryl oxygens in 3b has been ruled out as the ab initio optimization clearly shows that, at the global energy minimum, the closest distance from 2'-OH to any of the non-bridging phosphoryl oxygens is at least 3 – 4 Å.

The 2'-OH mediated intramolecular O–H···O H-bond is attached to the puckered sugar moiety in 3b and 4a, which expectedly causes the deviation from co-planarity, thereby making this non-linear H-bond rather weak in both nucleosides and nucleotides. At the low-energy minimum of our NMR constrained structures for both 3b and 4a, a rather long H-bonded bridge (~2.2Å with MM and ~2.0Å with ab initio, Table 3 in Paper IV) and considerably smaller /O-H…O3' bond angle over the temperature range studied [113.6° (288 K) ≤ /O-H…O3' ≤ 93° (368 K)] with MM and ~120° (298 K) with ab initio calculations, Table 3 in Paper IV).

5.1.2 Thermodynamics of inter- and intramolecular hydrogen bonding
The NMR lineshape analysis (Table 6 in Paper IV) of 2'-OH gave the ΔH_{298K}^{H-bond} of 7.5 kJ mol^{-1} for 3b and 8.4 kJ mol^{-1} for 4a; similar analyses of the methylene protons of 3'-ethylphosphate moiety in 3b also gave comparable ΔH_{298K}^{H-bond} of 7.3 kJ mol^{-1}. The donor nature of the 2'-OH in the intramolecular H-bonding in 4a is evident from its relatively reduced flexibility ([-T ΔS_{298K}^{‡}]_{2'-OH} = -17.9 (± 0.5) kJ mol^{-1}) because of the loss of conformational freedom owing to the intramolecular 2'O-H···O3' H-bonding, compared to the acceptor 3'-OH ([-T ΔS_{298K}^{‡}]_{3'-OH} = -19.8 (± 0.6) kJ mol^{-1}) at 298 K.

5.1.3 Nature of hydration around the 2'-hydroxyl group
The ROESY spectra for 3b and 4a at 308 K, in DMSO-d_6, show (Figure 6 in Paper IV) a clear positive rOe contact of 2'-OH of 3b and both 2'- and 3'-OH for 4a, respectively with water. The presence of hydrophilic 3'-phosphate group in 3b causes a much higher water activity in the vicinity of its 2'-OH, which in turn causes the 2'-OH to exchange faster (Figure 7 in Paper IV), culminating in a shorter exchange life-time (τ) for 2'-OH proton with HOD in 3b [τ_{2'-OH}: 489 ms] compared to that in 4a [τ_{2'-OH}: 6897 ms]. The activation energy (E_a) of the exchange (Table 6 in Paper IV) with the bound-water for 2'- and 3'-OH in 4a (48.3 and 45.0 kJ mol^{-1} respectively) is higher compared to that of 2'-OH in 3b [31.9 kJ mol^{-1}], thereby showing that the kinetic availability of hydrated 2'-OH in 3b for any inter and
intramolecular interactions, in general, is owing to the vicinal 3'-phosphate residue.

5.2 Implications

This study suggests that 2'-OH in native RNA can mediate other inter- or intramolecular interactions only in competition with the bound-water. The specific chemical nature of the vicinal 3' substituent, spatial orientation of 2'-OH (sugar pucker dependent, thereby dictating the proximity towards O3') and the H-bond angle (\(\angle\text{O-H}^-\text{O}\)) contribute to the strength of this type of weak non-linear H-bonding. The availability of the bound water around 2'-OH in RNA would, however, be dictated by whether the vicinal phosphate is exposed to the bulk water or not. This implies that relatively poor hydration around a specific 2'-OH across a polynucleotide chain, owing to some hydrophobic microenvironmental pocket around that hydroxyl in RNA, can make it more accessible to form H-bonding with other donor or acceptor functions.

6. Aromatic interactions and stacking

The intra- and intermolecular stacking and/or other aromatic interactions\(^\text{208-213}\) involving both non-biological (section 6.1.1) as well as biological systems (section 6.1.2) is a topic of fundamental interest related to molecular recognition and biological functionalities.

6.1 Aromatic interactions

The aromatic interactions can be categorized according to the geometries\(^\text{208,209,211}\) of interactions: face-to-face, edge-to-face (T-shaped) and offset (Figure 7). Several studies invoked that edge-to-face and offset stacked geometries are energetically favored over the face-to-face interaction among aromatic moieties. The major non-covalent aromatic interactions (mostly in non-biological model systems) so far identified can be
categorized as (i) $\pi - \pi$ interaction, (ii) CH – $\pi$ interaction (involving CH of both aryl and alkyl) and (iii) ion – $\pi$ interaction (involving both cation-$\pi$ and anion-$\pi$).

Hunter et al. invoked offset stacking (Figure 7b) involving attractive atom-$\pi\sigma$ interaction (electrostatic in nature) and edge-to-face interactions (Figure 7a) between two aromatic moieties, rather than repulsive $\pi-\pi$ interaction as proposed in the face-to-face stacking. However, recent studies invoked an attractive face-to-face aromatic interaction (Figure 7c) involving the negatively charged $\pi$ cloud of unsubstituted benzene ring and positively charged $\pi$ framework of the hexafluoro benzene ($C_6F_6$).

Theoretical studies recently showed that dispersion effects other than electrostatics dominate both aryl CH-$\pi$, and alkyl CH-$\pi$ interactions. However, in all cases alkyl CH-$\pi$ interactions are weaker than aryl CH-$\pi$ interactions. Dougherty et al. showed that electrostatic and polarization effects are the dominant contributions in the cation-$\pi$ interaction. However, the quadrupole moment and dispersion effect of aromatic system as well as charge transfer interaction occasionally play a secondary role in such processes. Nishio et al. proposed partial charge transfer arising from through-space proximity between alkyl hydrogen and aromatic moiety as the basis for CH-$\pi$ interaction (Figure 8B). On the other hand, Siegel et al., proposed a through-space polar (Coulombic)/$\pi$ contribution as a dominating factor in the electrostatic interactions involved in edge-to-face as well as the center-to-edge (i.e. offset) oriented aromatic moieties in the neutral 1,8-diarylnaphthalene system. Diedrich et al. also invoked similar through-space polar/$\pi$ contribution in the edge-to-face aromatic interactions. Moreover, such polar/$\pi$ contribution has also been showed in aromatic interactions in the ionic states involving carboxylate ion/arene and trimethylammonium ion/arene. However, Inoue et al. cited examples of ground state partial charge transfer process in intra- and intermolecular stacking involving indole and adeninium rings. Thus, the molecular nature of aromatic interactions is still now a major debate.

### 6.1.1 Aromatic interactions in non-biological model system

Crystal structure of the CT (1:1) complex between aniline and p-dinitrobenzene was the first structural report showing aromatic rings in stacked arrangement (Figure 8A). Stoddart et al. first identified the edge-to-face interaction in the collapsed empty cavities of crown ethers from solid-state structure. The direct experimental evidences of intramolecular stacking in solution state have come from temperature-dependent NMR studies of side chain substituted dibenzodiazocine derivatives, concentration-dependent $^1H$ NMR studies of bis-adenine with aliphatic linker, $^{syn/anti}$ epimerization by temperature-dependent NMR of 1,8-diarylnaphthalene and dynamic NMR studies of substituted benzyl...
pyridinium bromide. In all these studies the rotational free energy of model aromatic systems has been quantified by NMR, in order to show that the edge-to-face aromatic interaction and CH-π interaction are the driving force for the observed conformational isomerism. Further, the 1H NMR and solid-state studies of molecular zipper complex and metal tris-bipyridine complex,247 as well as evaluation of NH-π interaction driven intermolecular association by NMR241 showed the influence of such aromatic interaction in bimolecular complex formation. The geometrical dependence on the strength of aromatic interaction has been observed242 in flavoenzyme mimic. Recent studies228 invoked a weak non-covalent attractive anion-π interaction (Figure 8F) involving the anion and positively charged π framework of the hexafluoro benzene (C₆F₆). However, another studies229 showed an anion-arene interaction having both positive as well as negative components. The role of heteroatoms (i.e. polar component) and solvent effect in aromatic interaction have been studied217 by Gellman et al., which showed the predominance of polarizibility of aromatic surface over hydrophobicity as major attractive components in such interactions. On the other hand, Moore et al. studied243 aromatic interaction between m-phenylene ethynylene based cyclic and linear system showing the influence of hydrophobicity in such association. The solvent dependent aromatic interactions in the foldamers have been extensively studied244 by Iversson et al. The NMR studies with chemical double mutant cycle system as proposed by Hunter et al., showed209,245,246 the predominating electronic effect over dispersion forces in

---

**Figure 8.** Panel (A) – (F) shows the different types of aromatic interactions in non-biological model system.
edge-to-face aromatic interaction, which is, however, in contrast to what Wilcox et al. observed.219

6.1.2 Aromatic interactions in biological system
Evidences regarding the nature of intramolecular aromatic interactions in nucleic acids and their complexes214,248-267 have mainly come from various structural studies. Thus, Hunter et al. first invoked the presence of offset stacked nucleobases in DNA214 based on the X-ray crystallographic data followed by computer modelling to construct conformation-dependent energy maps based on van der Waals and electrostatic interactions calculated between stacked bases. Recent crystallographic studies in RNA have proposed the presence of stabilizing O4' (lone pair) – π (nucleobase),253 water (lone pair) – π (nucleobase)254 as well as water (H) – π (nucleobase)254 from the proximity calculations. Rooman et al.255,256,267 defined and analyzed stair-shaped motifs, which simultaneously involve base stacking, hydrogen bonding and cation-π interaction in protein-DNA complex through the geometrical proximity found in the X-ray crystallographic database. Recent database studies258 showed the importance of thymine-methyl/π interaction in the sequence-dependent deformability of DNA. Moreover, studies257 based on screening of nucleic acid databases showed that divalent cations [like Mg(OH)2]2+ interact favorably with π systems of nucleic acid bases. It has also been proposed257 that some critical cation-π interactions may contribute to the stability of the anticodon arm of yeast tRNAth and to the magnesium core of the Tetrahymena group I intron P4-P6 domain. Such cation-π interactions have been also implicated in DNA bending, DNA-protein recognition, base-flipping, RNA folding, and catalysis.257 Ab initio studies have shown the presence of the aromatic interactions (mostly of cation-π in nature)264 between protein and DNA involving positively charged Arg or Lys side chains and aromatic rings of nucleic acids. The X-ray studies along with calorimetric and fluorescence analyses have shown250-252 the importance of electrostatic cation-π interaction in the protein recognition of m7G part of the mRNA cap structure. Similar kinetic and calorimetric experiments260 have also identified the key aromatic π–π stacking interaction between Tyr41 and the adenine ring of bound nucleotides in the active site of an aminoglycoside phosphotransferase enzyme. The importance of edge-to-face aromatic interactions in preorganization of the peptide secondary structure has been studied by Waters et al.261-263

6.2 Base-Base stacking interaction in nucleic acids
Studies using vapour pressure osmometry,277 temperature and pH-dependent NMR,268-272,278 ORD274 as well as theoretical simulations273,275 have earlier been performed with nucleic acids to elucidate inter- and intramolecular stacking interactions. However, these studies have failed to show any insight
on the nature of such intramolecular base-base stacking. Recently, the dangling end stabilization studies\textsuperscript{279-283} have estimated the strength of stacking in both DNA and RNA duplex which has been found to be ca 0.4 – 3.6 kcal mol\textsuperscript{-1}. According to Kool et al.,\textsuperscript{284} aromatic stacking interaction between nucleobases in water involves electrostatics (dipole-dipole and dipole-induced dipole) interactions, dispersion (momentary dipole-induced dipole) effects and solvation. The stacking in single stranded nucleic acid is opposed by the conformational entropy\textsuperscript{2} since the rotational motion across the phosphate backbone and glycosyl (Scheme 1) in a dinucleotide step is rather unconstrained. However, earlier studies qualitatively showed the sequence dependency and negligible salt effect in single strand stacking. Very fast forward rate constant (10\textsuperscript{7} s\textsuperscript{-1}) for single strand stacking\textsuperscript{2} and single chemical shift (NMR time-average) at any pH validates (Papers VI - IX) the assumption for pH-dependent two-state model (stacking ⇌ destacking)\textsuperscript{271,272} in the NMR time scale.

6.3 \( pK_a \) perturbation in nucleic acids and protein folding

6.3.1 \( pK_a \) perturbation in nucleic acids

Determination and interpretation of \( pK_a \) of an ionizable group (a basic center for protonation or a dissociable group) in biomolecules highlights the molecular mechanism of its biological function. The shift in \( pK_a \) values is an important source of information about neighboring charges, electrostatics, structural perturbation as well as partial charge distribution over the whole molecule, and differential hydration of the microenvironment.

Structural studies have revealed that in a large number of RNAs (and in certain cases in DNAs),\textsuperscript{286-299} the \( pK_a \) values of nucleobases (particularly adenine and cytosine) are significantly perturbed relative to that of mononucleotides. Thus, the most effective acid-base catalysis can be performed with ionizable groups having \( pK_a \) near physiological pH of 7, thereby accounting for widespread use of histidines\textsuperscript{282} in protein enzyme active sites. NMR studies with leadzyme\textsuperscript{287,288} showed the perturbed \( pK_a \) of 6.5 for the adenine (A25) at the active site. Similarly, adenines within the internal loop at domain B of the hairpin ribozyme showed\textsuperscript{289} perturbed \( pK_a \) values ranging from 4.8 to 5.8. The shifted adenine \( pK_a \) values (6.2) for A\textsuperscript{+}C wobble base pair have been found\textsuperscript{290} in an internal loop in domain A of the hairpin ribozyme. A perturbed \( pK_a \) (~7.6) for an adenine (A2451) in the ribosomal peptidyl transferase centre has been reported.\textsuperscript{291} However, studies of Xiong et al.\textsuperscript{292} showed contradictory results, which proposed pH independent behavior of A2451 dimethylsulfate modification in ribozyme of \textit{Thermus aquaticus} and \textit{Mycobacterium smegmatis}. The perturbed \( pK_a \) (6.4) for uracil in the ternary Uracil DNA glycosylase complex has also been reported.\textsuperscript{293,294} Studies from Nakano et al.\textsuperscript{295} as well as Lupták et al.\textsuperscript{296} showed moderate perturbation of \( pK_a \) to near neutrality for active-site
cytosine (C75) in HDV ribozyme. Similar $pK_a$ perturbation for cytosine has also been observed\textsuperscript{297} from studies on self-cleaving ribozyme. Similarly, the A$^+$.C wobble base pairing involving adenine with perturbed $pK_a$ (6.6) has also been reported\textsuperscript{298,299} for DNA oligomer.

The shift of the $pK_a$ values of nucleobases at the active site in ribozyme towards neutral pH (in ground state rather than transition state perturbation\textsuperscript{173,300}) might arise from: (i) through the stabilization by H-bonding (by wobble base pairing\textsuperscript{287-289,298,299}), (ii) by salt-bridge formation\textsuperscript{288} or due to stabilizing (for the protonated nucleobase) and destabilizing (for the deprotonated nucleobase) interaction with the phosphodiester moiety in the steric proximity and (iii) due to change in structural microenvironment (by folding or unfolding).\textsuperscript{300}

Thus, it is likely that a local hydrophobic pocket around a specific nucleobase (created by folding) may reduce the local dielectric of the microenvironment, thereby increasing the $pK_a$ values\textsuperscript{128,300}. Alternatively, a relatively more exposure of a specific nucleobase to the aqueous medium may give a $pK_a$ value very similar to that of a mononucleotide counterpart. This is consistent with the fact that the $pK_a$ in general decreases (i.e. more acidic) as the dielectric of the medium increases.\textsuperscript{128} A particular example of this phenomenon can be found\textsuperscript{128} in the recent determination of $pK_a$ for –COOH group in salicyllic acid, which showed $pK_a$ of 6.6 in DMSO ($\varepsilon = 48$) and 2.9 in water ($\varepsilon = 78$).

However, the proton is sequestered by the H-bonding interaction through which its $pK_a$ perturbation occurs [i.e. by (i) and (ii)], thereby becomes ineffective in catalytic participation. On the other hand, the $pK_a$ perturbation via local microenvironment change [as described in (iii)] is considered to be more efficient way to steer the enzymatic properties (both in general acid-base catalysis\textsuperscript{170,173,183,300} and electrostatic catalysis\textsuperscript{173,285}) of RNA.

It has been commented\textsuperscript{300} that limited folding pattern (so far discovered), lesser diversity of polar side chains and negatively charged (hydrophilic) backbone in RNA compared to those of protein, apparently, restricts the ability of the former to modulate the microenvironment at the binding sites. On the contrary, the present pH-dependent study with ssRNAs and its comparative analyses with isosequential ssDNAs (section 7.3) has experimentally demonstrated, for the first time, that the the microenvironment around the nucleobases of RNA varies very widely, and consequently, their intrinsic pseudoaromatic characteristics changes in a variable manner depending upon their differential electrostatic modulation through sequence dependent nearest neighbor offset stacking interaction, which forms the current basis of the concept of the extended genetic code (see the implication part, section 7.4).

6.3.2 $pK_a$ perturbation in proteins
Unusual pKₐ values have been observed for many protein residues at the (i) active site and implicated in catalytic function, (ii) ligand-binding site and (iii) protein-protein interaction site. Such pKₐ perturbation may arise due to changing electrostatic properties of the binding surface, which modulates (a) charge-charge interaction between ionizable groups, (b) differential solvent accessibility of charged side chains and (c) H-bonding interactions. For example, a pKₐ 6.36 for glutamate side chain in rat CD2 has been attributed to mutual electrostatic repulsion. A pKₐ of 7.5 for aspartate side chain in oxidized *Escherichia coli* thioredoxin has been explained in terms of hydrophobic local environment. A perturbed pKₐ at the binding site of Nicotinic Acetylcholine receptor has recently been shown. The pKₐ of histidine residues at the C-terminus and N-terminus of the folded α-helices of barnase protein are shown to be ~0.5 unit higher and 0.8 unit lower, respectively, compared to the unfolded residue. This perturbation of pKₐ results from the combination of the different electrostatic environments. The increase of pKₐ at C-terminus is due to the charge-helix dipole interaction and stabilization through H-bonding whereas the lowering of pKₐ has been attributed to the movement of side chain away from protein thereby promoting solvent induced electrostatic screening. The positioning of a charged group affects its pKₐ in protein as shown in the active site Lys of acetoacetate decarboxylase, in which its pKₐ is perturbed by ~4 units to a value of 6 due to presence of a nearby Lys residue.

Recent NMR titration studies with turkey ovomucoid third domain (OMTKY3), and DNA binding proteins from archaeum *Sulfolobus solfataricus* (Sso7d) showed that pKₐ of a particular residue was sensed by resonances not only in that residue but also from neighboring residue. In OMTKY3, the major interactions responsible for lowering of pKₐ of P₁-Glu19 are H-bonding between carboxylate of P₁-Glu19 and hydroxyl group of P₂-Thr17, the charge–charge interaction between P₁-Glu19 and P₃-Arg21, and the intra-residue H-bonding between carboxylate of P₁-Glu19 and its own amide group. However, the elevated pKₐ of P₂5’-Glu43 might arise due to helix-dipole interaction. The perturbation of pKₐ’s of the different side chains in Sso7d are due to the H-bonding (between Glu35 and Tyr33 residues), electrostatic screening as well as formation of salt bridge (between Asp34 and Tys20 residues) within the active site. Thus, the local electrostatics mediated pKₐ perturbation has been shown to contribute site-specifically to the pH dependence of the protein stability. Several other studies have come up showing the importance of the electrostatic interaction of neighboring residues in the overall stability and functions of proteins.

6.4 Biological importance of single stranded nucleic acids
Single dangling nucleotide at the end of both DNA and RNA duplexes is known to increase the duplex stability. In more recent studies, it has been shown that longer single-stranded dangling residues (up to tetranucleotide) stabilize the RNA-RNA and DNA-DNA duplexes even slightly more (by an extra ≈0.1 – 1.0 kcal mol⁻¹) than the single-nucleotide dangling end (≈2.0 kcal mol⁻¹). The ssRNA as dangling end stabilizes the helix, which results into specific biological function: (i) The CCA overhang at the 3'-terminal of tRNA, which is involved in aminocacylation reaction, also stabilize the cloverleaf structure of tRNA. (ii) The dangling end adjacent to the codon-anticodon basepair stabilize the interaction between mRNA and tRNA. (iii) The dangling nucleotide at the 3'-end of a pseudoknot RNA is also known to stabilize the stem structure. (iv) Single unpaired base bulges in RNA duplexes enhance the stability of the RNA more compared to the fully base paired counterpart, in which both the base identity as well as the nearest neighbor context have been shown to be important for the overall stability of the bulges. (v) Recognition and interaction with many ligands including proteins also take place with ssRNA.

The comparison of ssDNA dangling-end with that of ssRNA showed that the ssDNA motif with 5'-dangling ends or ssRNA with 3'-dangling ends contributes to stability equally or more than their ssRNA or ssDNA counterparts respectively. Rosemeyer and Seela have also shown that polarisibility of dangling base correlates with the stability of duplex better than the hydrophobicity.

Many ssDNAs show their functional properties upon binding to specific proteins. These ssDNA binding proteins belong to either of the two following categories: (i) those that recognize a particular sequence of nucleic acids like transcriptional regulation, telomere replication and (ii) those that specifically interact with a particular physical form of nucleic acids in a sequence independent manner like *Escherichia coli* ssDNA binding protein, RepA protein or RecA protein. The stacking interaction between aromatic side chain of protein and nucleobases is considered to be a major force in the binding process for those proteins in the second category, which bind to the single-stranded nucleic acids (SSB proteins). Stacking between aromatic amino acids and nucleic acid bases (mainly enthalpic in nature) plays an important role in the enzyme specificity with nucleic acid substrate.

It has long been qualitatively known that the ssDNA exists with some ordered single-stranded structure. Recently, ssDNA templates have been shown to direct multistep small molecule synthesis programmed by DNA sequences, thereby showing the possibility of transposing the ssDNA chirality (natural or engineered) to the product of the DNA-templated synthesis. An in-depth understanding of the preorganized conformation of
the ssDNA as well as ssRNA is clearly a pre-requisite to understand their various intra and intermolecular functionalities.

7. Present Work (Papers V – X)

7.1. Tandem nearest-neighbor aromatic interactions in nicotinamide derivative (Paper V)

Yamada et al. has demonstrated\textsuperscript{332} that by selective shielding of one side of the pyridinium face by the intramolecular face-to-face stacking (by X-ray crystallography) of the neighboring phenyl ring in a nicotinamide derivative (methyl derivative of 8a), nucleophiles attacks only from the non-shielded side to give exclusively 1,4-adduct over the 1,6-adduct in 99\% ee.

\textbf{Scheme 7.} Compounds used in NMR and \textit{ab initio} studies (\textbf{Paper V}). The small arrow indicates possible rotational torsions. The p\textsubscript{K\texttextsubscript{a}} values assigned to each proton derived from Hill plots (Figure 2 and Table 1 in \textbf{Paper V}).

The pH-dependent \textsuperscript{1}H NMR studies showed that the basicity of the pyridinyl group (p\textsubscript{K\texttextsubscript{a}} \textasciitilde 2.9 – 3.0) in 8a \textrightarrow 8a\textsuperscript{+} (Scheme 7) could be measured both from the pH-dependent (1 \textless pH \textless 7) chemical shifts (Figure 1 in \textbf{Paper V}) of the pyridinyl protons as well as from the protons of the neighboring phenyl and methyl groups. Thus the pyridinium moiety in 8a (p\textsubscript{K\texttextsubscript{a}} \textasciitilde 2.9 – 3.0) becomes more basic compared to that in the standard 9a (p\textsubscript{K\texttextsubscript{a}} 2.56) due to this nearest-neighbor electrostatic interaction between phenyl and pyridinium moieties.
7.1.2 Nearest-neighbor interaction between pyridinyl and phenyl groups

The differences in relative shielding of pyridinyl protons in protonated $8a^+$ compared to that in $9a^+$, with reference to their neutral counterparts $8a$ and $9a$ respectively (section A in Results and Discussion part in Paper V) demonstrates that the H5/H9-edge of the pyridinyl group is more affected than the H7/H8-edge (Figure 9) from the neighboring phenyl ring, thereby showing a direct evidence of the edge-to-face cation (pyridinium)-π (phenyl) interaction.

The quicker relaxation of pyridinyl (pyridinium) protons in $8a$ or $8a^+$ compared to that of the standards $9a$ and $9a^+$ shows that former relax through the protons of the neighboring phenyl group respectively. The $\Delta T_1$ shows (Table 1 in Paper V) that, for both $8a$ and $8a^+$, the H5 relaxes more quickly than any other protons of the pyridinyl ring [$\Delta pK_a \sim 0.5$ from $\delta H5$, which is slightly (~0.1) higher than all other pyridinyl protons] followed by H9 and the slowest relaxing protons being H7 and H8. This supports the preferential interaction of H5/H9 edge of pyridinyl group with the π face of the phenyl moiety. Thus, this $T_1$ relaxation study clearly backs up the proposed CH (pyridinyl)-π (phenyl) interaction at the N-state, and cation (pyridinium) – π (phenyl) interaction in the P-state.

The ab initio geometry optimization based molecular modeling also supported the edge-to-face interaction in that the distance between phenyl- and H5 proton of pyridine/pyridinium groups of $8a$ and $8a^+$ decreases by 0.7 Å (Table 4 in Paper V) upon protonation, thereby substantiating our observation (see section D in the Result and Discussion part in Paper V) that an electrostatic pyridinium (cation)-phenyl (π) interaction is relatively stronger than the neutral CH (pyridinyl)–π (phenyl) interaction.

7.1.3 Cascade of pyridinium-phenyl-methyl cross-talk

The fact that the pH-dependent chemical shifts (Figure 1 in Paper V) of the phenyl and methyl protons gives the $pK_a$ of the pyridine moiety of $8a$ (but not the methyl protons of $9a$) also suggests that the nearest neighbor cation (pyridinium) -π (phenyl) interaction also steers the CH (methyl) -π (phenyl) interaction in tandem. It therefore shows that the whole pyridine-phenyl-methyl system in $8a$ is electronically coupled at the ground state, cross-modulating the physicochemical property of the next neighbor. This cation (pyridinium)-π (phenyl) interaction is indeed more stable ($\Delta \Delta G_{pK_a}^{\circ} = -2.1 \text{ kJ mol}^{-1}$) than that of the CH (methyl)-π (phenyl) interaction ($\Delta \Delta G_{pK_a}^{\circ} = -0.8 \text{ kJ mol}^{-1}$).

7.2 Implications

The present pH-dependent $^1$H NMR study has given a straightforward experimental evidence of the intramolecular cation-π interaction in edge-to-

face geometry thereby perturbing the $pK_a$ of the pyridine-nitrogen in presence of neighboring aromatic group. Thus this can be used as a major experimental tool to identify the nature of aromatic interaction in a system having a protonation or a deprotonation site. Interestingly, such titration method also gives an enormous insight into the energetics, geometry as well as the nature of weak non-covalent contribution in such aromatic interactions, which are of great importance in molecular recognition pattern in both biological as well as non-biological system.

7.3 Cross-modulation of physicochemical character of nucleobases in single stranded nucleic acids (Papers VI – X)

Specific recognition of the single-stranded nucleic acids is a fundamental requirement in most of the important biological processes like telomere recognition, DNA replication and repair, transcription, translation and RNA processing (see section 6.4). In absence of base pairing, the stacking interaction plays a more important role in the self-assembly of both ssDNA and ssRNA structures

7.3.1 pH-dependent titration of dinucleotides and other ssRNAs and ssDNAs (Papers VI – IX)

The dinucleotides 10a – f (Scheme 8) for these studies are chosen such that only one of the two nucleobases in the molecule can be exclusively protonated or deprotonated at a given pH (at 298 K) in order to show the effect of alteration of the electronic character of one aglycon on the other. The pH titration by $^1$H-NMR for each nucleobase in 10a – f shows (Figures 1 and 2 as well as Table 1 in Paper VI) that $pK_a$ of a certain nucleobase can be obtained from not only its own aromatic maker protons, but also from the

![Scheme 8](image)

B = 9-Quaninyl (G); 9-Adeninyl (A);
1-Cytosinyl (C); 1-Uracilyl (U); 1-thyminyl (T)
Sequence Specific Nearest-neighbor Interaction In Dinucleotides

\[ \Delta \delta (\text{in ppm}) = \delta_{\text{neutral}} - \delta_{\text{protonated}} \] for Protonation and \[ \delta_{\text{neutral}} - \delta_{\text{deprotonated}} \] for deprotonation

Figure 10
marker protons of the neighboring nucleobase (to the best of our knowledge these are the first report in nucleic acids, however there are such evidences in protein, see section 6.3.2) as a result of cross-modulation of two-coupled \( \pi \) systems of neighboring aglycones. This also shows the sequence-specific effect in two possible isomeric aglycone combination in purine-pyrimidine (10c and 10d), purine-purine (10a and 10b) as well as pyrimidine-pyrimidine (10e and 10f) dimers. Similarly, pH-dependent titration (at 298 K) using \(^1\text{H} \) NMR has been performed for all tri-, tetra-, penta- and hexameric ssRNAs and ssDNAs (12a – 15c, Scheme 8) for which the sequences are chosen in such a way that over the whole titration range only single base ionization effect (at 5’G) can be observed along the strand.

7.3.2 Sequence specific nearest-neighbor interaction and thermodynamics of the offset stacking (Papers VI and IX)

The pH-dependent chemical shift change (\( \Delta \delta \), Figure 10) in any of the aromatic marker protons in either of the two coupled nucleobases in dinucleotide monophosphates (10a – f) shows variable electrostatic modulation depending upon the geometry of the offset stacking, partial charge of the heteroatom as well as the sequence. The thermodynamics of the pH-dependent offset stacking in dinucleotides (10a – f) are shown in Figure 4D in Paper VI. Among all these dimers, in CpU\(^-\) the formation of 1-uracilylate as a function of pH failed to promote any destacking \( \Delta \delta_{N-D} > 0 \), see Figure 10 and \([\Delta G_{\text{stacking}}]_{N\text{-state}}: -2.3 \text{ kJ mol}^{-1} \) compared to \([\Delta G_{\text{stacking}}]_{D\text{-state}}: -1.9 \text{ kJ mol}^{-1} \). In general, the extent of stacking decreases (Figure 4 in Paper VI) in the following order: 5’-Purine-Purine-3’ \( \cong \) 5’-purine-pyrimidine-3’ > 5’-pyrimidine-pyrimidine-3’ > 5’-Pyrimidine-Purine-3’ at the N-state. Similarly the pairwise comparison of the \( \Delta \delta_{(N - D)} \) of all aromatic marker protons (Figure 6 in Paper IX) in the isosequential ssDNAs and ssRNAs (10b, 10g, 11a – 15c, Scheme 8) shows how the edge of the nucleobase responds to the interplay of the nearest neighbor electrostatic interactions (see section 7.3.6) due to 9-guanylate formation. The \( \delta \)H8A in ssRNA is relatively more destacked \( \Delta \delta_{(N - D)} < 0 \), showing its significant pH-dependent change \( i.e. \) greater modulation), compared to that of \( \delta \)H2A. On the other hand, this trend is reversed in ssDNA, where \( \delta \)H2dA changes more as a function of pH compared to that of \( \delta \)H8dA, thereby showing the different stacking geometry in ssDNA compared to ssRNA as further supported by the NMR constrained MD simulation (see section 7.3.7). It was first argued that the basis of this variable \( \Delta \delta \) is the partial charge transfer (see ref. 27 in Paper VI) between two coupled nucleobases, however, later with further studies on single stranded trimers and oligomers (Papers VII – IX), electrostatics has been proposed (sections 7.3.4 and 7.3.5) as the mechanistic basis for such cross-modulation. The distinction between partial charge transfer and electrostatics is delicate as well as debatable although the
later mechanism is more likely as neither of the dimers or oligomers (Scheme 8) showed any characteristic UV absorption in the ionic state.

7.3.3 $pK_a$ shift of 9-guanylate due to electrostatic effect of 3’- and 5’-phosphate versus nearest-neighbor nucleobase (Papers VI – IX)

Comparison (Table 1 in Paper VI) of the $pK_a$ of 9-guaninyl in GpEt (3a) (9.25 ±0.02) and EtpG (6a) (9.57 ±0.01) shows that the 5’-phosphate makes the $pK_a$ of 9-guaninyl more basic ($[\Delta pK_a]_{6a} - (3b) \approx 0.32$) compared to the 3’-phosphate. This is due to the electrostatic interaction as a result of the spatial proximity of the negatively charged 5’-phosphate and the fused imidazole moiety of the 9-guaninyl (in the anti conformation) thereby enhancing the electron-density in the fused pyrimidine moiety giving an overall increase of $pK_a$ of 9-guaninyl. The pairwise comparison at the dimer and trimer level, however, shows (Table 1 in Papers VI, VII and IX): in GpA (10b, 9.17 ±0.02)/ GpC (10g, 9.56 ±0.01) giving $[\Delta pK_a]_{10g} - (10b) \approx 0.41$; however, in deoxy series, d(GpC) (11b) is only slightly more basic than d(GpA) (11a), giving $[\Delta pK_a]_{11b} - (11a) \approx 0.1$; similarly, in ApG (10a, 9.71 ±0.02)/ CpG (10h, 9.45 ±0.01) giving $[\Delta pK_a]_{10a} - (10b) \approx 0.26$ as well as in GpApA (12a, 9.75 ±0.02)/ GpApC (12b, 9.88 ±0.01) showing $[\Delta pK_a]_{12b} - (12a) \approx 0.13$. Thus, the $pK_a$ of N1-H of 9-guaninyl residues is sequence dependent since the two isomeric dimers and/or trimers have the same phosphate charge but different 3’-nucleobase. Moreover, $[\Delta pK_a]_{10a} - (10b) \approx 0.54$ is relatively high compared to $[\Delta pK_a]_{10g} - (10b) \approx 0.11$ showing that nearest-neighbor modulation in pG is opposed by the 5’-phosphate effect when compared to Gp in similar dimeric sequences. This suggests that the chemical nature of the nucleobase steers the $pK_a$ of the nearest neighbor nucleobase(s) more effectively than the phosphates although latter being negatively charged also imparts electrostatic effect through phosphate-nucleobase interaction. These studies (Paper IX) also demonstrated that the 9-guaninyl moieties in ssDNA are relatively more basic (calculated from $\delta$H8G) than those in the corresponding ssRNAs ($[\Delta pK_a]_{ssDNA-ssRNA} \approx 0.17$ for hexamer, 0.29 for pentamer, and 0.11 for tetramer, as found in monomers $[\Delta pK_a]_{dGpEt-GpEt} \approx 0.15$) with the exception of d(GpC) and d(GpApC).

7.3.4 Variation of $pK_a$ of 9-guaninyl among different marker protons across the single strand (Papers VII – IX)

The $pK_a$ of 5’-Gp residue (Table 1 in Paper VII) from its own $\delta$H8G varies from 9.76±0.01 (in 14a) to 9.88±0.01 (in 12b) among oligo-ssRNAs 12b and 14a – c and that from 9.64±0.02 (in 13a) to 10.11±0.01 (in 15b) among oligo-ssDNAs 13a and 15a – c. The variation of $pK_a$ ( $\Delta pK_a$) for 5’-Gp residue as measured from the other aromatic marker protons of various nucleotide residues (Table 1 in Paper VII) across the strand in ssRNA and ssDNA: $\Delta pK_a \approx 0.9$ and 0.2 for hexameric 14c and 15c respectively, $\Delta pK_a \approx 0.5$ and 0.14 for pentameric 14b and 15b respectively. Similar $\Delta pK_a$ for tri-
and tetrameric ssRNA and ssDNA are relatively small. Thus, the corresponding \( \Delta G_{\text{pK}_a}^\circ \) values (in kJ mol\(^{-1}\)) of \(-5.1\) and \(1.1\) for 14c and 15c respectively as well as that of 2.9 and 0.8 for 14b and 15b, respectively, have been attributed to the variable strength of electrostatic interactions across the strand in ssRNA compared to ssDNA. This considerably less intramolecular nucleobase-dependent modulation of p\(\text{K}_a\) of 9-guaninyl in ssDNA compared to the isosequential ssRNA suggests that the pseudoaromatic character of the nucleobases in ssDNA is much less tunable (mainly Type 1 effect, in which the p\(\text{K}_a\) of 9-guaninyl from the neighbor is the same as that of the ion source, 9-guanylate ion) than those in ssRNA (mainly Type 2 effect, in which the p\(\text{K}_a\) of 9-guaninyl from the neighbor shows to be larger than that of the ion source, 9-guanylate ion, owing to its own intrinsic electronic property). Studies with proteins (section 6.3.2) have predominantly showed the perturbation with lowering of p\(\text{K}_a\), however, few studies have also reported elevated p\(\text{K}_a\).\(^{308-310}\) Thus, our experimental evidences of Type 2 effect (elevated p\(\text{K}_a\) of 9-guaninyl from the marker protons of neighboring nucleobases in a sequence context) have shown new aspects of local electrostatic modulation via nearest neighbor interactions in single stranded nucleic acids.

7.3.5 Propagation of electrostatic interplay across the single strand (Papers VIII and IX)

It has been observed that the propagation of the interplay of various electrostatic forces (section 7.2.6) across the ssDNA chain is considerably less favourable (NMR detectable effectively up to the fourth nucleobase moiety in hexameric ssDNA 15c) than what we observed for the corresponding isosequential ssRNA 14c (NMR detectable up to the sixth nucleobase residue, upto \(-21\) Å in the unfolded state). This suggests that the intramolecular attractive offset stacking is stronger in ssDNA than those of ssRNA as the cost of stacking (i.e. creation of hydrophobic environment) is high for ssRNA compared to ssDNA owing to the presence of hydrophilic 2'-OH group in the former. Thus the stacking is more easily perturbed in ssRNA, upon generation of a repulsive electrostatic charge (i.e. 9-guanylate ion) compared to ssDNA.

7.3.6 ssDNA is better stacked than that of ssRNA (Papers VIII and IX)

The different magnitude of the upfield shifts (i.e. differential anisotropic shielding) of different marker protons shows how the edge of a particular marker proton feels the aromatic ring current of the nearest neighbor. The evidences supporting a stronger stacked structure for ssDNA strand compared to the ssRNA come from two independent observations: (i) the aromatic marker protons in ssDNA are more anisotropically shielded compared to those in the isosequential ssRNA [\(\Delta \delta_{\text{(deoxy-ribo)}}\), Figure 4 in Paper IX], (ii) they also have larger upfield oligomerization shift [Figure 5
in *Paper IX*] with respect to the monomeric counterparts compared to those of isosequential ssRNA and (iii) The larger negative $\Delta \delta_{(N-D)}$ of aromatic marker protons (except the 3'-terminal pC in 14b and 14c) in ssRNAs compared to those for ssDNAs (Figure 6 in *Paper IX*) shows relatively more pronounced destacking of the neighboring nucleobases in the former owing to the 9-guanylate formation.

The various interplay of repulsive anion(G')-$\pi$/dipole(Im$^{-\delta}$) and counteractive atom-$\pi\sigma$ and anion(G')-$\pi$/dipole(Py$^{+\delta}$) [Paper VIII] electrostatic interactions [which have been further dissected and elaborated in details in section B(ii) in *Paper IX*] contributes to the relative stacking propensity across the single strand. This interplay (Scheme 2 in *Paper IX*) for GpA$^1$p step in ssRNA compared to d(GpA$^1$p) step in ssDNA has been proposed on the basis of the pairwise comparison of $\Delta \delta_{(N-D)}$ (Figure 6 in *Paper IX*). It shows that the overall attractive component of this electrostatic interplay (as a function of pH) in ssDNA series is stronger than that of isosequential ssRNA counterparts *i.e.* more repulsive in the latter compared to the former [section B(v) in *Paper IX*]. Thus, the electrostatic interplay at GpA$^1$p step (both in ssRNA and ssDNA) will, therefore, modulate the pseudoaromaticity of pA$^2$p, which will further change the electrostatic interaction between pA$^2$p and pA$^3$p thereby modulating the pseudoaromatic character of pA$^3$p as well as propagating this destacking process (depending upon the strength of these interactions) across the helix (Scheme 2 in *Paper IX*). Thus, the destacking across ssDNA is less, thereby giving stronger stacking, compare to the corresponding ssRNA counterpart.

7.3.7 Tentative stacking geometry in single stranded nucleic acid from NMR constrained MD simulation: a qualitative approach (Paper IX)

The NMR-MD derived geometry of hexameric ssRNA (14c) with C3'-endo sugar geometry and ssDNA (15c) with C2'-endo sugar geometry are shown to be like A-type and B-type helix (Figure 8 in *Paper IX*), respectively. Thus, the nucleobase overlap at each dinucleotide step of the 15c and 14c (Figures 9C and 9D in *Paper IX*) shows that the electron-rich imidazole stacks above the electron-deficient pyrimidine in 5'→3' direction ([5'-imidazolyl moiety]→[3'-pyrimidinyl moiety] electrostatic interaction) in 15c while, in contradistinction, the pyrimidine stacks above the imidazole in 5'→3' direction ([5'-pyrimidinyl moiety]→[3'-imidazolyl moiety electrostatic interaction]) in 14c. This is fully consistent with the model of the electrostatic interplay proposed [section B(ii) in *Paper IX*].

7.3.8 Relative dissection of stacking vis-à-vis base pairing from the $pK_a$ calculation of the model mononucleotides (Paper X)

The pH-dependent $^1$H NMR showed the $pK_a$s of both nucleoside 3'-ethylphosphates (3a – 3e for ribo and 3f – 3j for deoxy, Scheme 8) as well as nucleoside 3',5'-bis-ethylphosphates (5a – 5e for ribo and 5f – 5j for deoxy,
Scheme 8) in both 2'-deoxy and ribo series under an uniform NMR condition. The nucleoside 3',5'-bis-ethylphosphates (5a – 5e for ribo and 5f – 5j for deoxy, Scheme 8) can mimic the internucleotidic monomeric unit of DNA and RNA in which the stacking contribution is completely absent. All monomeric DNA nucleobases are found to be more basic than the corresponding RNA nucleobases by ~0.2 pKₐ unit. The ΔpKₐ values for the G-C basepairing with model ribo pair (5a/5c) and deoxy pair (5f/5h) are 5.04 and 5.24 respectively (Table 2A in Paper X). Similarly, the ΔpKₐ for the A-U/T basepairing with model ribo pair (5b/5d) and deoxy pair (5g/5i) are 5.53 and 6.29 respectively. Comparison of these ΔpKₐ values of G-C and A-T/U basepairing shows that the ribo basepairing is stronger (i.e. ΔpKₐ is less) than the corresponding deoxy basepairing. The ΔpKₐ is considerably less for the G-C basepairing in both ribo and deoxy series than those for the A-U/T basepairing (Table 2A in Paper X) which is consistent with the fact that the former is stronger than the latter.¹²

The comparison of linear plots for each of [ΔG°_{bp}]_{RR–DD} and [ΔG°_{stacking}]_{RR–DD} as a function of % A-T/U bp content (Figure 3 in Paper X) with opposite slope shows that with the increasing content of A-T/U base pairs the stability of DNA-DNA duplex weakens over the corresponding RNA-RNA duplexes ([ΔG°_{bp}]_{RR–DD}), while the strength of stacking ([ΔG°_{stacking}]_{RR–DD}) of A-T rich DNA-DNA sequence increases in comparison with A-U rich sequence in RNA-RNA duplexes. This increased stacking contribution from T compared to U, in DNA-DNA over RNA-RNA duplex, comes from favorable electrostatic CH/π interaction²⁵⁸ between the 5-methyl group of T with the nearest-neighbor A in the AT rich sequence.

7.4 Implications

The net result of this electrostatic cross-talk between two neighboring aglycones as a result of base-base stacking is creation of a unique set of aglycones in an oligo or polynucleotide, whose physico-chemical properties are completely dependent upon the nearest neighbor electrostatic interactions. This has considerable implication in the specific ligand binding ability, aptamer recognition, RNA catalysis, and most probably in codon-anticodon interaction. The poorer pseudoaromatic cross-modulation and stronger stacking in ssDNA compared to ssRNA (as manifested in their respective pKₐ's) may have considerable implications in serving the purpose of DNA as a carrier of the genetic code (its almost error-free replication and transcription). On the other hand, the conformational flexibility of ssRNA allows it to create different scaffolds and nascent folded states with the characteristic tunable dielectrics, giving variable microenvironments (resulting in to larger pKₐ variation), which is manifested in its dual biological role in general, as we witness in the translation machinery and catalysis.
Thus in a RNA sequence, $P_1^1Q_1^1N^1Q_2^2P_2^2$, the actual physico-chemical integrity of $N$ is dictated by the pseudoaromatic character of both neighboring $Q_1$ and $Q_2$, whose properties are further tuned by the electronic nature of $P_1$ and $P_2$. Hence, the relative stacking ⇪ destacking in any two adjacent nucleotides will actually set the ON and OFF switch for the tunability of the pseudoaromatic character of a particular nucleobase, $N$. Thus, the pseudoaromatic character of $N$ can have at least $2^4$ numbers of variations, depending upon the chemical nature of the neighboring $Q_1$ and $Q_2$, which therefore implies that a given nucleobase sequence in a polynucleotide chain constitutes an unique extended genetic code, which can be turned ON or OFF depending upon the intrinsic dynamics of folding and unfolding within the molecule owing to the sequence context or interaction with an external ligand.
8. Acknowledgement

This work was carried out at the Department of Bioorganic Chemistry, Biomedical Centre, Uppsala during the years 1998 – 2003.

A lot of faces that flashes across my mind now! My gratitude is beyond the words scribbled here.

To begin with, I would like to express my sincere appreciation and gratitude to my supervisor Professor Jyoti Chattopadhyaya for giving me this opportunity to work in his well-equipped laboratory, for introducing me to the field of structural and biophysical aspects of nucleic acids, for his thoughtful guidance, support and constructive criticism throughout these years. Though it was a tough time to be a Ph. D student, every day was a great experience!

I wish to express my deep gratitude to my all colleagues throughout these years in the department for their company and friendship: Subhrangsu Chatterjee, S. Karthick Babu, Oommen P. Varghese, Dimitri Honcharenko, Dimitri Ossipov, Peter Agback, Melker Holmer, Anil Kumar, Mrinal Kanti Kundu, Xiaoguang Luo, Tatiana Maltseva, Jan Milecki, Catherine Petit, Anna Trifonova, Irina Velikian, Pei Wen and Edouard Zamaratski.

My special thanks to:
Pradeep Kumar, for your great friendship and those off-hour socio-political discussions. Johan Isaksson, for being always helpful as a colleague with smiling face and for those interesting chat during "helium filling". Jharna Barman and Pradeep Cheruku, for taking the pain in preparation and purification of many of these compounds used in our studies. Oleksandr Plashkevych for all your prompt help with computer related problems. Andras Földesi, for being so cooperative in everything. Christophe Thibaudeau, for showing me the intricate details of NMR, PSEUROT, UNIX and many more things during my very first year in this department. Anders Sandström - a nice personality. I miss those discussions (starting from computer to Swedish and French societies) in every morning coffee break during my first few weeks in this department. Nitin Puri, being so helpful during my initial year in Uppsala. Devapriya Choudhuri, for helping me installing Gaussian and for rare brand of adda.

Jaana Öst, Pernilla Brohage and Rabab Elkarib for their kind assistance and competent secretarial work.
All staffs from BMC-biblioteket and from BMC administration for their kind help.

I am grateful to my teachers during school and university days. In particular, I would like to remember with gratitude: Dr. Sovan Chatterjee (Ballygunge Govt. High school, Calcutta, India), who had first instigated my interest in chemistry. Professor R. R. Rao and Professor Aloka Rao (Visva Bharati, Santiniketan, India), who introduced me with the fascinating aspects of organic chemistry. I still remember your first lecture, which depicted "why we should be proud of being an organic chemist". Though in this age of interfacial science, the distinctive subject-wise barrier is rapidly vanishing, it was a major boost for a fresher! Professor Asish De (Indian Association for the Cultivation of Science, India), who has amazing humane warmth in his personality. Thank you for all your help.

I would like to thank all my friends in India (the list will be too long to mention in person) for being friends. My warm regards to the ever-growing Indian community in Uppsala and Stockholm, which never lets one feel away from home by arranging stimulating parties and get-together. Thanks in particular to Bharath Bhikkaji (we started our innings in Uppsala more or less at the same time), for being a nice friend and for sharing similar enthusiasm about cricket. Thanks to Robert and gang of the Uppsala cricket team for providing me with the opportunity to refresh my cricket passion, even in Sweden.

I would like to express my special gratitude to Biplab & Suparna Sanyal; Santanu Dasgupta & Lena Åslund; Sabyasachi & Somali Sanyal, for being affectionate family friends.

I have no words to express my feelings for my family back home in India, my Father (from whom I first heard the word "research" in my childhood, and got inspired seeing his life long pursuance and endeavor in academics) and my Mother (who taught me the rare human quality "patience", which is like a light still showing me the path and I always feel proud of you, Ma) for your endless love, care and support. I also like to thank my other relatives and my in-laws for being so supportive.

My sincere thanks to my wife Sandipta, for being not only my colleague but also my best friend in life; thank you for all your love, care and affections as well as for bearing with my irritably odd (late hours) lab schedule.
9. References

44

47


189. Williams, D. J.; Boots, J. L.; Hall, K. B. *RNA*. **2001**, 7, 44.


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 Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology. (Prior to October, 1993, the series was published under the title “Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science”.)