UPPSALA
UNIVERSITET

Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology 164

Some Aspects of Physicochemical Properties of DNA and RNA

SANDIPTA ACHARYA

ACTA
UNIVERSITATIS
UPSALIENSIS
UPPSALA
2006

ISSN 1651-6214
ISBN 91-554-6518-8
urn:nbn:se:uu:diva-6741

Dissertation presented at Uppsala University to be publicly examined in Room C10301 (floor 3), Biomedical Center, Uppsala, Saturday, April 29, 2006 at 12:00 for the degree of Doctor of Philosophy. The examination will be conducted in English.

Abstract
Acharya, S. 2006. Some Aspects of Physicochemical Properties of DNA and RNA. Acta Universitatis Upsaliensis. *Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology* 164. 74 pp. Uppsala. ISBN 91-554-6518-8.

This thesis is based on nine research publications (**I – IX**) on structure and reactivity of RNA vis-à-vis DNA. The DNA and RNA are made of flexible pentose sugar units, polyelectrolytic phosphodiester backbone, and heterocyclic nucleobases. DNA stores our genetic code, whereas RNA is involved both in protein biosynthesis and catalysis. Various ligand-binding and recognition properties of DNA/RNA are mediated through inter- and intra-molecular H-bonding and stacking interactions, beside hydration, van der Waal and London dispersion forces. In this work the pH dependant chemical shift, pK_a values of 2'-OH group as well as those the nucleobases in different sequence context, alkaline hydrolysis of the internucleotidic phosphodiester bonds and analysis of NOESY footprints along with NMR constrained molecular dynamics simulation were used as tools to explore and understand the physico-chemical behavior of various nucleic acid sequences, and the forces involved in their self-assembly process. **Papers I – II** showed that the ionization of 2'-OH group is nucleobase-dependant. **Paper III** showed that the chemical characters of internucleotidic phosphate are non-identical in RNA compared to that of DNA. **Papers IV – VI** show that variable intramolecular electrostatic interactions between electronically coupled nearest neighbor nucleobases in a ssRNA can modulate their respective pseudoaromatic character, and result in creation of a unique set of aglycons with unique properties depending on propensity and geometry of nearest neighbor interaction. **Paper VII** showed that the cross-modulation of the pseudoaromatic character of nucleobases by the nearest neighbor is sequence-dependant in nature in oligonucleotides. **Paper VIII** showed that the purine-rich hexameric ssDNA and ssRNA retain the right-handed helical structure (B-type in ssDNA and A-type in ssRNA) in the single-stranded form even in absence of intermolecular hydrogen bonding. The directionality of stacking geometry however differs in ssDNA compared to ssRNA. In ssDNA the relatively electron-rich imidazole stacks above the electron-deficient pyrimidine in the 5' to 3' direction, in contradistinction, the pyrimidine stacks above the imidazole in the 5' to 3' direction in ssRNA. **Paper IX** showed that the pK_a values of the nucleobases in monomeric nucleotides can be used to show that a RNA-RNA duplex is more stable than a DNA-DNA duplex. The dissection of the relative strength of base-pairing and stacking showed that the relative contribution of former compared to that of the latter in an RNA-RNA over the corresponding DNA-DNA duplexes decreases with the increasing content of A-T/U base pairs in a sequence.

Keywords: nucleic acids, hydrogen-bonding, stacking, single-stranded, NMR, molecular dynamics

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

ISSN 1651-6214
ISBN 91-554-6518-8
urn:nbn:se:uu:diva-6741 (http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-6741)

To The Memory of My Father

THE ORIGINAL PUBLICATIONS

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

I. Velikyan, I.; **Acharya**, **S**.; Trifonova, A.; Földesi, A. and Chattopadhyaya, J. The pK_a's of 2'-Hydroxyl Group in Nucleosides and Nucleotides.
J. Am. Chem. Soc. **2001**, *123*, 2893-2894.
First authorship is shared by Velikyan, I. and Acharya, S.

II. **Acharya**, **S**.; Földesi, A.; and Chattopadhyaya, J. The pK_a of the Internucleotidic 2'-Hydroxyl Group in Diribonucleoside (3'→5') Monophosphates.
J. Org. Chem. **2003**, *68*, 1906-1910.

III. Barman, J.; **Acharya, S.**; Chuanzheng, Z.; Chatterjee, S.; Engström, A. and Chattopadhyaya J. Non-identical electronic characters of the internucleotidic pohosphates in RNA modulate the chemical reactivity of the phosphodiester bonds.
Org. Biomol. Chem. **2006, *4*,** 928-941.
First authorship is shared by Barman, J. and Acharya, S.

IV. **Acharya**, **S**.; Acharya, P.; Földesi, A. and Chattopadhyaya, J. Cross-Modulation of Physicochemical Character of Aglycones in Dinucleoside (3'→5') Monophosphates by the Nearest Neighbor Interaction in the Stacked State.
J. Am. Chem. Soc. **2002**, *124*, 13722-13730.

V. Acharya, P.; **Acharya**, **S**.; Földesi, A. and Chattopadhyaya, J. Tandem Electrostatic Effect from the First to the Third Aglycon in the Trimeric RNA Owing to the Nearest-Neighbor Interaction.
J. Am. Chem. Soc. **2003**, *125*, 2094-2100.

VI. Acharya, P.; **Acharya**, **S**.; Cheruku, P.; Amirkhanov, N. V.; Földesi, A. and Chattopadhyaya, J. Cross-Modulation of the pK_a of Nucleobases in a Single-Stranded Hexameric-RNA Due to Tandem Electrostatic Nearest-Neighbor Interactions.
J. Am. Chem. Soc. **2003**, 125, 9948-9961.

VII. **Acharya**, **S**.; Barman, J.; Cheruku, P.; Chatterjee, S.; Acharya, P.; Isaksson, J. and Chattopadhyaya, J. Significant pK_a Perturbation of Nucleobases Is an Intrinsic Property of the Sequence Context in DNA and RNA.
J. Am. Chem. Soc. **2004**, 126, 8674-8681.

VIII. Isaksson, J.; **Acharya**, **S**.; Barman, J.; Cheruku, P. and Chattopadhyaya, J. Single-Stranded Adenine-Rich DNA and RNA Retain Structural Characteristics of Their Respective Double-Stranded Conformations and Show Directional Differences in Stacking Pattern.
Biochemistry **2004**, 43, 15996-16010.

IX. Acharya, P.; Cheruku, P.; Chatterjee, S.; **Acharya**, **S**. and Chattopadhyaya J. Measurement of Nucleobase pK_a Values in Model Mononucleotides Shows RNA-RNA Duplexes To Be More Stable than DNA-DNA Duplexes.
J. Am. Chem. Soc. **2004,** 126, 2862-2869.

Contents

Abbreviations

B	Nucleobase
D	Deprotonated state
DD duplex	DNA-DNA duplex
DNA	DeoxyriboNucleic Acid
EF-Tu	Elongation Factor Tu
HPLC	High Performance Liquid Chromatography
H-bond	Hydrogen bond
N	Neutral state
N-type	North type
NMR	Nuclear Magnetic Resonance
MD	Molecular Dynamics
mRNA	Messenger RNA
NOESY	Nuclear Overhauser Effect Spectroscopy
P	Protonated state
PAGE	Polyacrylamide Gel Electrophoresis
R	Pearson Correlation Coefficient
RNA	Ribonucleic Acid
RNAi	RNA interference
rRNA	Ribosomal RNA
RR duplex	RNA-RNA duplex
S-type	South type
ss	single-stranded
tRNA	Transfer RNA
ΔG	Free energy of a process
ΔH	Enthalpy of a process
ΔS	Entropy of a process

1. Physicochemical properties of Nucleic acids

1.1 Structure of Nucleic acids

Nucleic acids are important bio-molecules endowed with cellular functions like conservation, replication, and transmission of genetic information, recognition as well as catalysis. Based on 2'-substitution, nucleic acids can be classified into two types *i.e* (i) deoxyribonucleic acids (DNA) and (ii) ribonucleic acids (RNA). A varied number of modifications (about 93) in nucleobase and the 2' substituent[1,2] are however present in nucleoside units of tRNA, rRNA and mRNA. DNA and RNA polymers are built of nucleoside monomer units bound to each other through 3'→5' phosphodiester linkage[2-4]. Each nucleoside unit is made up of a five membered D-pentofuranose sugar unit connected to a heterocyclic nucleobase (purine or pyrimidine) through a N-gycosidic linkage[2-4] (Figure 1 panel A) The sugar unit is either β-D-2'-deoxyribosyl or β-D-ribosyl in case of DNA or RNA respectively. The 1' carbon of 2'-deoxyribose in case of DNA is connected to either of the nucleobases adenine at N_9, guanine at N_9, cytosine at N_1 or thymine at N_1[2-4].

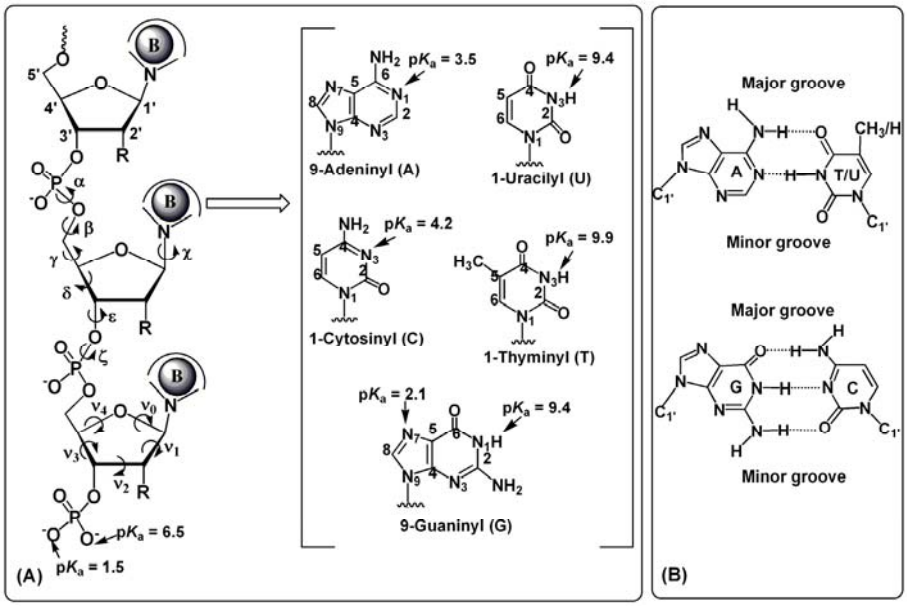


Figure 1. Panel (A) shows the endocyclic (ν_0-ν_4) torsions of pentafuranose and sugar-phosphate backbone torsions ($\alpha,\beta,\gamma,\delta,\varepsilon$ and ζ)[2]; constituent nucleobases and

phosphates for DNA and RNA with respective pK_a values of ionization sites (with small arrows). Panel (B) shows Watson-Crick basepairing found in usual double stranded DNA and RNA. Each double helix has a major groove and a minor groove; minor groove is on the side of the base pair where the sugar is attached.

In RNA the nucleobase thymine is replaced by uracil, while all other bases remain unaltered (Figure 1 panel A).

The torsion angles α, β, γ, δ, ε and ζ define the conformation along the sugar-phosphate backbone in nucleic acids[2-4] (Figure 1 panel A). For the pentose ring of both the ribose (in RNA) and deoxyribose (in DNA), the five endocyclic torsion angles are specified as ν_0, ν_1, ν_2, ν_3 and ν_4[2-4] (Figure 1 panel A). A planar five-membered ring is sterically and energetically very unfavorable[5,6]. To relieve the strain the puckered forms interconvert continuously through a pseudorotational cycle[2-5,7-12] (Figure 2). The puckered geometry of the pentofuranose ring in nucleic acids can be described by two parameters[8-12] (i) phase angle P (indicating which part of the ring is mostly

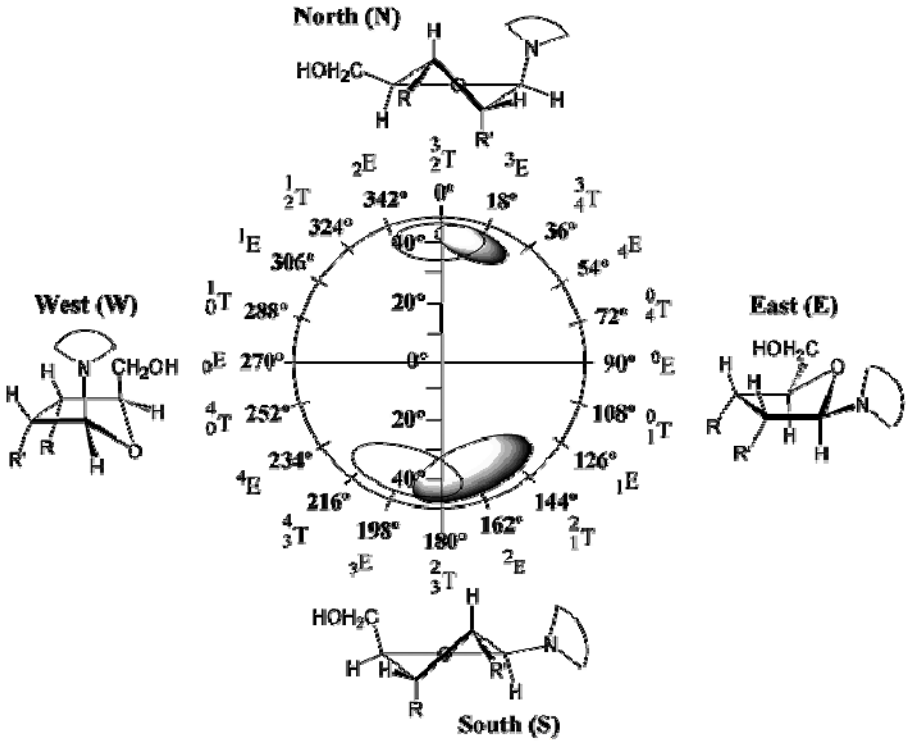


Figure 2. The pseudorotation wheel (E = envelope; T = twist) for pentafuranosyl D-nucleosides. The hyperspace of geometries accessible to N-and S-type pseudorotamers is within the shaded circle ($-1° < P_N < 34°$, $137° < P_S < 194°$, $30° < \psi_m < 46°$) for β- D-nucleosides and within unshaded circle ($-18° < P_N < 19°$, $168° < P_S < 224°$, $28° < \psi_m < 49°$) for α- D-nucleosides.

puckered) (ii) puckering amplitude ψ_m (indicating the largest deviation of the endocyclic torsion from zero). The crystal structures of nucleos(t)ides[13] suggest that the conformation of the pentofuranose can be adequately described by a two-state North (N, C3'-*endo*-C2'-*exo*) ⇄ South (S, C2'-*endo*-C3'-*exo*) equilibrium model (energy barrier of N- and S-type pseudorotational equilibrium is 1.2-5 kcal/mol)[5,14], since no other state is found to exist abundantly. Only a few East-type[15] (E, O4'-*endo*) pseudorotamers and no West-type (W, O4'-*exo*) conformers were found among the crystal structures[13]. W-type conformers are energetically disfavored owing to the pseudoaxial orientation of both the nucleobase and the 5'-CH_2OH group as well as the eclipsed C2' and C3' substituents. Analogously, the energy destabilization of E-type conformations, compared to either N- or S-type can be attributed to the eclipsed orientation of the C2' and C3' substituents. Solution phase NMR studies[16-21] also showed that the confromation of pentafuranose can be described by a two-state North ⇄ South equilibrium. This has been observed in some B ⇄ Z DNA[16,17], A ⇄ Z RNA[18,19] and A-form ⇄ B-form lariat RNA[20,21] transitions. Nature of nucleobase[5,22,23], sugar modification[24,25] as well as 2'-and/or 3'-substitution[5,26,27] can be used to drive the sugar conformation towards either North-type or South-type.

The torsion angle about the glycosidic bond is specified by the angle χ[2-4] (Figure 1 panel A). The two ranges found for the χ value are designated as syn ($-90^o \leq \chi \leq 90^o$) and anti ($90^o \leq \chi \leq 270^o$). In syn conformation the six-membered ring of the purines or the carbonyl at C2 of pyrimidines is near to the sugar, which makes this conformation unfavorable. The anti conformation is energetically more favourable with the six-membered ring of the purines or the carbonyl at C2 of pyrimidines away from the sugar unit.

In duplex nucleic acids, nucleobases present in one strand of the polymeric nucleic acid chain engage in base pairing with the complementary nucleobases of the opposite strand with the help of H-bonding[28-33] and stacking among the neighboring nucleobases of the same strand[34-39]. Thus base pairing and stacking are the predominant forces stabilizing the secondary structures of nucleic acids. Hydration[40] of phosphate backbone furanose oxygen and nucleobases as well as positioning of the phosphate backbone towards the exterior of the double helix to minimize repulsive forces are other factors stabilizing nuclei acid structure. Nucleic acids generally form Watson-Crick type base pairing[28] where nucleobases adenine (A) and thymine (T) (in DNA)/uracil (U) (in RNA) [designated as A**.**T/U] or guanine (G) and cytosine (C) [designated as G**.**C] are involved in H-bonding (Figure 1 panel B). Non-Watson-Crick type base-pairing (non-canonical base-pairing) (Figure 3) like Hoogsteen (A**.**T/U) [2,3,41,42], Wobble (G**.**U, G**.**T) [2,3,43-45], reverse Hoogsteen (A**.**C, A**.**U) [3], reverse Wobble (G**.**U) [3] to name a few, exists in addition to Watson-Crick in the vast array of nucleic acid self-assemblies. In many instances ionization of bases can provide further opportunities for base pairing as in triplex formation ($(C.G)C^+$)[3,46]

DNA is generally found in the duplex form but single stranded (see section 6) and circular forms[47,48], hairpins[49,50] triplexes[3,46] and quadruplexes[51] also exist. RNA generally functions in the single-stranded form and organizes (folds itself) to secondary and tertiary structures[52] like hairpins, bulges, loops, pseudoknots, through stacking and hydrogen bonding across the folded motifs according to structural and functional requirements. Depending on the amount of hydration and the nature of counter-ions present the DNA and RNA duplex can organize itself forming A, B or Z forms[53-61] which are different in their helical parameters. B form is the predominant form of DNA duplex. Z form of DNA is made up of (pyrimidine-purine)$_n$ tandem repeats with alternating conformation (anti and syn for glycosylic torsion and South and North for sugar conformation) for glycosyl torsion and sugar [56-61]. Earlier sequences with alternating (CG)$_n$ repeats were known to

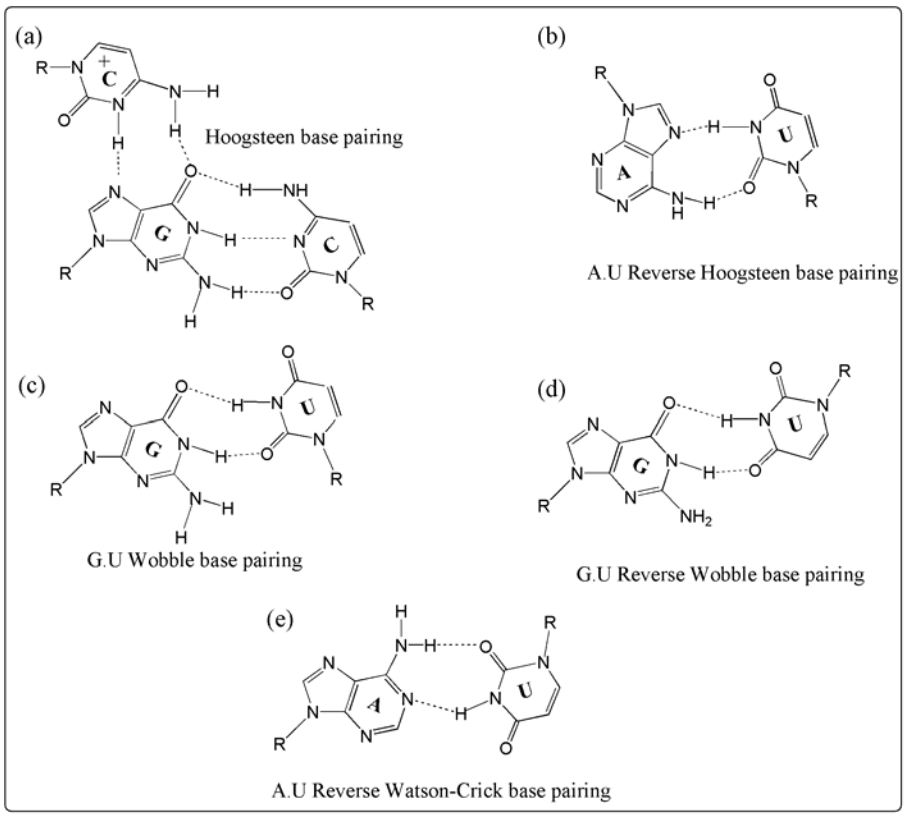


Figure 3. Panel a-e shows some of the different kinds of non-canonical base pairing involved in DNA and RNA secondary structures.

be responsible for Z-DNA formation but other repetitive sequences[59-61] of (AC)$_n$ and (TG)$_n$ are known now. RNA on the other hand exists predominantly as the A type helix. It has been found that the pentose sugar confor-

mations in B-form DNA helix (*i.e* deoxyribose sugar) is C2'-*endo i.e* S-type, and that of A-form RNA helix (*i.e* ribose sugar) is C3'-*endo i.e* N-type. The detailed features of A, B and Z forms can be found in Table 3 in section 6.3.

The arrangement of bases in the base paired oligonucleotide, as well as their stacking mode can be specified by a set of structural parameters[62]. They are defined in terms of translation and rotation about the coordinate axis. The translational parameters can be classified into a) translation involving two bases of a base pair: X displacement, Y displacement, stagger, stretch and shear and b) translation involving two successive base-pairs: rise, slide and shift. The rotational parameters can be classified into c) rotation involving two bases of a base pair: tip, inclination, opening, propeller twist and buckle and d) rotation involving two successive base-pairs: twist, roll and tilt.

1.2 Reactive groups in Nucleic acids

A detailed study of the structural aspects of nucleic acid helps us to understand the mechanisms behind its reactivity and function. The groups that impart reactivity to nucleic acids can be broadly classified as follows: (1) The *nucleobase* moieties help to structurally preorganize the strands by nearest-neighbor stacking[34-39] interactions and by H-bonding[28-33] to give the thermodynamically stable duplexes. It also assists in binding to ligands as in the aptamers[63-65] (*aptamers are nucleic acids that can bind with high affinity and specificity to a wide range of ligands*). (2) The *2'-OH group* in RNA, on other hand, is a quintessential function (see section 2.1) involved in all key biological transesterification reactions such as splicing[66], RNA catalysis[66] or base catalysis by RNA cleaving proteins[67]. (3) Finally, some specific *internucleotidic phosphates in RNA*, which are the cleavage points (with or without metal ion cofactors)[66] for the biological transesterification reactions. It is indeed a very complex task to identify whether the chemical characters of all adeninyls or guaninyls or cytosinyls or uracilyls are similar or different compared to their respective monomeric counterparts when placed in a large sequence-context. Due to the fragile nature of the phosphodiester bonds in RNA it is also very difficult to probe if the chemical nature of all internucleotidic phosphate groups or 2'-OH groups are dissimilar when placed in a real biological context in various folded states with or without cofactors (such as metal ions or proteins that typically serve to stabilize the transition states for the cleavage reaction)[66].

Fortunately, nucleic acids have a variety of ionization centres present in the nucleobases, 2'-OH and phosphate groups throughout the pH range of 1-14 (Figure 1, panel A)[68]. Each of the nucleobase has a protonation or a deprotonation site in the acidic or the alkaline pH respectively, which itself could be a measure to understand, for example, if the aromatic characters of all 9-adeninyl moieties are similar or dissimilar depending upon the se-

quence context. The 9-adeninyl moiety in adenosine has a protonation site at N1with a pK_a value of 3.5. The 9-guaninyl moiety in guanosine has a protonation site at N7 with a pK_a value of 1.6 and a deprotonation site at N1 with a pK_a value of 9.2. Both uridine and thymidine have a deprotonation site at N3 with pK_a values 9.2 and 9.7 respectively. Cytidine has a protonation site at N3 with a pK_a of 4.2. In case of RNA, the 2'-OH group has an ionization site with pK_a around 12.0-14.0 (section 2.3 as well as **Papers I** and **II** in this thesis). Internucleotide phosphate groups have an ionization pK_a of 1.5[69] whereas the terminal phosphate has pK_as of 1.5 and 6.5 [69,70]. This means that at the physiological pH the internucleotide phosphates are fully ionized. Thus a careful pH titration of an oligo-DNA or -RNA allows us to probe the protonation or deprotonation equilibrium (pK_a) of the nucleobases to assess their aromatic characters as the sequence-dependant environment changes, and also to assess its effect on the neighboring nucleobases (see section 4.4), phosphates (see section 3.3) as well as 2'-OH (see section 3.3). Thus the pH-dependant protonation or deprotonation of nucleobases can be used as a tool to gain a deeper understanding of electronic properties and reactivity of nucleic acids, which in turn may help us to understand recognition and interaction processes in nucleic acids, in general. Similarly, the nucleobase-dependant ionization of 2'-OH groups in nucleotides can also be used to determine how the unique character of a specific nucleobase can alter or modulate the pK_a of the 2'-OH, in a variable manner thereby affecting its ability to participate in the transesterification reaction with the vicinal phosphate.

1.3 Forces underlying the stacking interactions of nucleobases in DNA and RNA

In nucleic acids, stacking of nucleobases allow DNA and RNA to form self-assembled structures. Thermodynamic data on single-stranded[71,72] and duplex oligonucleotides[73,74] show that stacking is driven by a favourable enthalpy factor organizing a helix. In duplexes this preorganization helps in the formation of H-bonding with its complementary strand. The nucleobases in an oligonucleotide constitute heterocyclic aromatic moieties with a relatively positively charged σ framework sandwiched between a π electron cloud (see section 4.1). It is to be also noted that the purine nuclobases even in its neutral state have inequal charge distribution with the imidazole ring being more electron rich than pyrimidine part[75]. In a oligomeric DNA or RNA, neighboring nucleobases in a stacked conformation can interact with each other due to the inherent polarity in them (the dipole moment values for each nucleobase is as follows: adenine 2.56D, guanine 6.55D, cytidine 6.39D, thymine 4.31D and uridine 4.37D)[31]. The negatively charged π cloud of one

ring can interact with the positive σ framework of the nearest-neighbor nucleobase and result in attractive interactions (see section 4.1). Similarly if either the π cloud or the σ fromework of two neighboring aromatic rings come close they repel each other[76.] Development of additional charge or creation of differential charge distribution in the system (by neighboring nucleobases and phosphate as well as formation of hydrophobic pockets) can increase or decrease the stability of the stacked nucleobases in nucleic acid single strands as well as duplexes.

2. Reactivity of the 2'-hydroxyl group in RNA

2.1 Importance of the 2'-hydroxyl group in RNA

The 2'-OH is the functional group that differentiates RNA from DNA. This group influences pentofuranose conformation and helix geometry[77,78], coordinates metal ions[79], provides a scaffold for solvent or protein interactions[80], and mediates catalysis[66,67,81-93] as well as tertiary interactions[94-115] by hydrogen bonding in RNA. The importance of the presence of 2'-OH group of RNA can be classified into (i) recognition[80,116-125] by RNA binding macromolecules like ribozymes[116-121] and proteins[122-125], (ii) processing and catalytic[66,67,81-93] properties of RNA as well as (iii) stabilization of RNA tertiary structure[94-115].

2.1.1 Role of the 2'-hydroxyl group in recognition

In large RNAs the 2'-OH group has a key role in substrate recognition during the formation of the enzyme-substrate complex as in the case of RNase P RNA[116,117], the Group I self splicing intron[118-122] and polymerase ribozyme[123]. Several 2'-OH contacts help the RNase P ribozyme[117] (to recognize the tertiary structure of pre-tRNAs), Group I[118-122] (to recognize the helix that present the 5' splice site by using four discrete 2'-OH) and RNA dependant RNA polymerase (to recognize the substrate with the help of eight 2'-OH groups present in the substrate)[123] to specifically recognize the substrate. Other than ribozymes, proteins also employ 2'-OH groups for sequence independant RNA recognition. Many proteins, binding double stranded RNA (dsRNA), contain a sequence called the dsRNA-binding motif (dsRBM) that depends mostly on 2'-OH contacts[80,124]. The specific interaction of RNA-activated protein kinase with dsRNA involves molecular recognition of a network of 2'-OH groups [125].

2.1.2 Role of the 2'-hydroxyl group in processing and catalytic properties of RNA

The catalytic RNA molecules–the ribozymes can be classified into i) self-splicing introns (Group I[81-90,66] and Group II introns[83-90,66]) and RNase P[83-90, 106,107] RNA which undergo RNA processing, ii) small self-cleaving ribozymes[83-90] like Hammerhead, Hairpin, HDV and Neurospora VS and iii)

the ribosome[83]. Reactions at phosphate center in ribozyme take place in two ways: (1) Ribonuclease P and self-splicing introns (Group I and Group II) catalyse phosphodiester-cleavage and ligation reactions that produce 5'-phosphate and 3'-hydroxyl termini (a, c and d in Figure 4). (2) The small self-cleaving ribozymes catalyse reversible phosphodiester cleavage reactions that generate 5'-hydroxyl and 2'–3'-cyclic-phosphate termini (b in Fig-

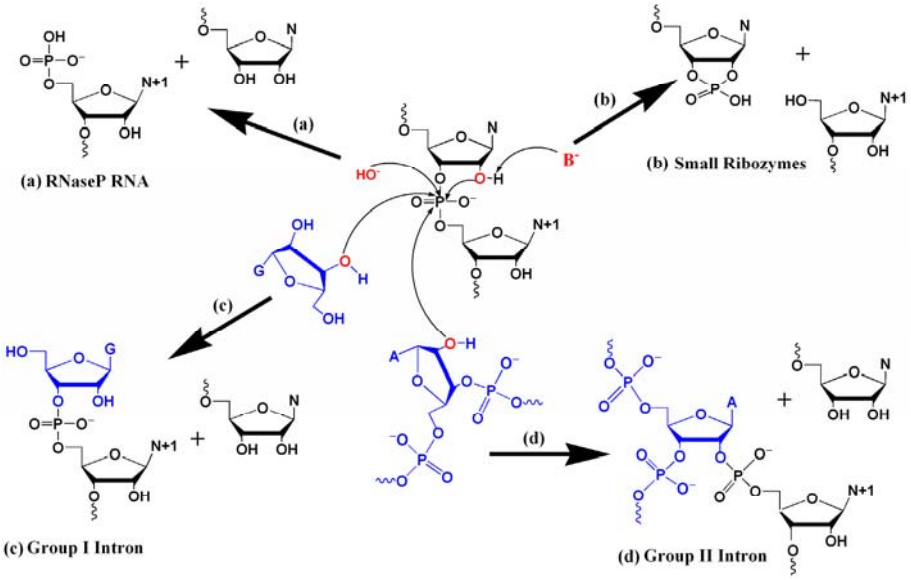


Figure 4. The different mechanistic pathways followed by different ribozymes in catalysis. The nucleophile used and the products formed in the phosphodiester cleavage reaction via transesterification differ with ribozyme type showing the diversity in catalytic activity. (a) In the case of RNase P RNA the nucleophile is a free water hydroxide (**HO^-**) or metal-bound hydroxide ion, which promotes the hydrolysis with the formation of 3'-OH and 5'-phosphate residues. (b) In small ribozymes hydrated metal ions (in hammerhead and HDV) or nucleobases (adenosine or cytosine in Hairpin and HDV), denoted as **B^-**, deprotonate the 2'-OH of the ribose attached to the nucleobase N in the figure and attacks the vicinal phosphate group leading to the formation of cyclic 2'-3' phosphate and 5'-OH group at the N+1 base. (c) The Group I intron splicing takes place by the attack of the 3'-OH (ionized with the help of solvated metal ions) of the exogenous guanosine cofactor (blue) at the phosphorus of the 5' splice site. The guanosine becomes covalently linked to the 5'-end of intron and a new 3'-OH is formed at the 3'end of exon. (d) The Group II intron splicing takes place by the attack of the 2'-OH group of a adenosine residue located within the RNA intron (blue) at the phosphrous of the 5' splice site. Upon self splicing the adenosine forms the lariat RNA by getting covalently linked to the 5'-end of intron through 2'-5' linkage with the N+1 base (first base of intron sequence) and a new 3'-OH is formed at the 3'end of exon. An S_N2 type inline attack mechanism is followed in all ribozymes. A pentacoordinated phosphorane intermediate is formed at the phosphorus center undergoing nucleophilic attack with inversion of configuration.

ure 4). Both kinds of phosphoryl-transfer reaction proceed via transesterification through S_N2 like mechanism[83,84,87,91] (Figure 4) with an in line attack of a nucleophile (HO^-) at the scissile phosphate, followed by the departure of a leaving group. This involves an inversion of configuration of the nonbridging oxygens of penta-coordinated phosphorane[91] (Figure 4). The nucleophilic attack at the scissile phosphorus center of the phosphodiester backbone in the RNA substrate is made by an oxyanion from a remote 2'-OH in Group II introns, the 3'-OH of exogeneous guanosine in Group I introns (ionized with the help of solvated Mg^{2+}), or by a free water hydroxide (HO^-) or Mg^{2+} bound hydroxide ion in RNase P. In small ribozymes like Hammerhead, Hairpin, HDV and Neurospora VS the nucleophilic attack on the reactive phosphate, on the other hand, is made by the neighboring 2'-OH. Metal ions like solvated Mg^{2+} (in case of Group I, RNaseP, hammerhead and HDV ribozymes)[79] and nucleobases with environmentally perturbed pK_a[126-129] values (in case of hairpin[126,127] and HDV[128,129] ribozymes) have however also been found to play prominent role in RNA catalysis. Recently it has also been shown that ribosomal RNA catalyses peptide bond synthesis in the ribosome[83] where 2'-OH also plays an important role. Apart from ribozymes RNA strand scission has been shown to be catalysed by ribonucleases like RNase A[67,92,93] which also involve nucleophilic attack of the 2'-OH on adjacent phosphrous atom.

2.1.3 Role of 2'-hydroxyl group in stabilization of RNA tertiary structure

The 2'-hydroxyl groups of RNA molecules often play important roles in RNA tertiary structure formation, both as hydrogen bond donor and acceptors, and in some cases to co-ordinate structurally important metal ions [66,94]. Various motifs like the U-turns[110,111], tetraloops[106-109], ribose zipper motif[96], A-minor motifs[101] and tetraloop-helix interactions[95,100,111] use hydrogen bonding abilities of 2'-OH group to stabilize tertiary structures of RNA. Studies on the structure of $tRNA^{phe}$ [97-99] show that the 2'-OH moieties of non-helical nucleotides are involved in hydrogen bonds to the nitrogenous bases, phosphate oxygen and other ribose groups. 2'-OH groups present in the minor groove of RNA duplexes are involved in the formation of a ribose zipper motif[96,102] as found in the crystal structure of the P4-P6 domain of Group I intron, HDV[104,110] and hairpin ribozymes[104,110]. Similar to ribose zipper motifs, A-minor motifs formed from hydrogen bonded 2'-OH groups are important for RNA tertiary structure stabilization in ribosomal RNA[101]. In small RNA hairpin loops the ribose hydroxyls participate in loop tertiary structure[106-109,112] stabilization as in UUCG tetraloop[106,108,109,112] the GNRA[107] (N stands for any nucleobase and R stands for purine) tetraloop as has been established by biochemical and NMR studies.

2'-OH groups are involved in structural stabilization of RNA-protein[114] and aptamer[115] complexes as well. Recent studies show the evidence of participation of the 2'-OH group in $tRNA^{phe}$ of *Thermus thermophilus* and $tRNA^{cys}$ of *E. coli* to stabilize its complex with EF-Tu[113] and cysteine-tRNA synthetase[114] respectively. The exchange properties of 2'-OH of a guanosine residue involved in a novel H-bond has been shown to contribute to the immobilization of bound AMP by the RNA aptamer[115].

2.2 Variability in the experimentally determined pK_a values of 2'-OH group of RNA

The biological importance of 2'-OH group in terms of understanding the structure and function of RNA in molecular details made the determination of pK_a of 2'-OH an important issue. It is however impossible to measure the pK_a of 2'-OH in a large RNA accurately as it is decomposed under alkaline conditions.

Variable pK_a values of 2'-OH[130-142] have been reported in the past for nucleos(t)ides[130-139] and internucleotidic 2'-OH in diribonucleoside 3'→5' monophosphates[140,141] as well as in an oligo-DNA with a single diribonucleoside (3'→5') monophosphate unit incorporated within[142]. In nucleosides and nucleotides they provide inconsistent values of the pK_a of 2'-OH group for the same compound mainly because of employment of different techniques like thermometric titration[130], electrometric titration[131], potentiometric titration[133] and quantum chemical calculations[133] to obtain the pK_a value. On the other hand studies on pK_a values of internucleotidic 2'-OH in diribonucleoside (3'→5') monophosphate[140,141] as well as a single diribonucleoside (3'→5') monophosphate unit embedded in a DNA oligomer[142] have been performed in different temperature and salt concentration conditions, using different techniques like HPLC[141] and PAGE[140,142] analysis for separation and quantification of the reaction components for determination of the pH-dependant first-order rate constants for the alkaline hydrolysis. It is well known that for every 10K change in temperature the pK_a values for acids and bases differ approximately by 0.1 to 0.3 units[143,144]. For a change in K^+ ion concentration from 0.5 to 3.0M the pK_a of the 2'-OH group change by 0.6 pK_a units[142]. Thus due to varied experimental conditions employed for 2'-OH pK_a calculation, the values already available in literature could not be compared directly.

2.3 Present work (Papers I – II)

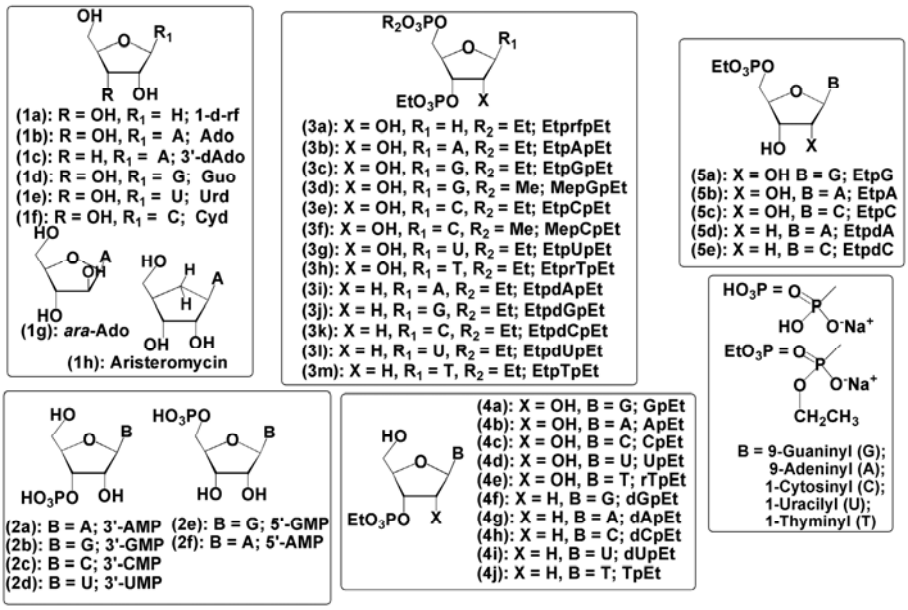
The present work showed how the nature of nucleobase and 3'-substituent can affect the pK_a of 2'-OH group in RNA. This has been shown by measurement of pH dependant ^{1}H shifts (for mononucleos(t)ides) as well as the variation of pseudo first order rate constant with pH (for dinucleotides).

2.3.1 Determination of pK_a of the 2'-OH group in nucleosides, mono- and dinucleotides and 3'→5' monophosphates by pH titration studies

The experimental pK_a values for 2'-OH were measured for mono and dinucleotide units[145]. The 2'-OH pK_a values of mononucleotidic compounds were obtained from pH dependant sugar proton (H1', H2', H3') chemical shifts of ribonucleosides (**1b**, **1d**, **1e**, **1f,** Scheme 1), their 3', 5'-bis-alkyl phosphodiester derivatives (**3b**, **3c**, **3d**, **3f,** Scheme 1), 3'-monophosphates (**2a**, **2b**, **2c**, **2d,** Scheme 1), adenosine 3'-ethyl phosphate (**4b,** Scheme 1), 3'-deoxyadenosine (**1c,** Scheme 1), *ara*-adenosine (**1g,** Scheme 1), aristeromycin (**1h,** Scheme 1), the abasic sugar (**1a,** Scheme 1**)** as well as abasic 3',5'-bis-ethyl phosphate (**3a,** Scheme 1) under identical condition. For dinucleotidic compounds (**6a – h** in Scheme 2) a plot of pH versus pseudo first order mean rate constant (k_{mean}) for alkaline cleavage was used to calculate the pK_a of the internucleotidic 2'-OH. The mean rate constant (k_{mean}) used was calculated from the area of proton (^{1}H) signals. As in these dimeric compounds (**6a-h**) in many cases the H2' and H3' chemical shifts could not be extracted due to overlapping of signals, the pH dependant chemical shift change was not used to calculate the pK_a values of internucleotidic 2'-OH.

2.3.2 Variation in pK_a values of 2'-OH group in nucleosides and mononucleotides with varying 3' substituents and aglycones

The pH-dependant proton chemical shift measurements under identical conditions for nucleosides and nucleotides (**1a-h, 2a-d, 3a-d**, **3f** and **4a,** Scheme 1) showed a pK_a variation of 2'-OH of upto 1.9 units (12.15-14.05) with maximum standard error of ±0.08 pK_a unit. From detailed studies of pK_a's for 2'-OH in various nucleosides and nucleotides, it is evident that electron withdrawing groups in the vicinity of 2'-oxyanion delocalizes the negative charge and stabilizes the system, resulting in increased acidity of the 2'-OH group. Effects which can cause delocalization of the oxyanion negative charge are i) H-bonding, ii) through-space field effect, iii) through-bond inductive effect, iv) solvation, or v) sterioelectronic anomeric and *gauche* effects. Thus the pK_a of the 2'-OH in adenosine (**1b**) is more acidic by 0.87 pK_a units compared to 3'-deoxy adenosine (**1c**). The presence of H-bonding between 2'-OH and 3'-OH with 2'-OH acting as donor in the former while its



Scheme 1

absence in latter could be one of the reason for increased acidity of 2'-OH in adenosine. Relative stabilization of the 2'-oxyanion by vicinal 3'-substituent has been shown by comparing the pK_a's of 2'-OH in 3'-AMP (**2a**) (13.81), adenosine 3'-ethyl phosphate (**4b**) (13.00), adenosine (**1b**) (12.15) and 3'-deoxyadenosine (**1c**) (13.02). The stabilization by 3'-substituent follows the order 3'-OH > 3'-H ≈3'-OPO_2EtO^- > 3'-OPO_3^{2-}. Apart from 3'-substituents, the chemical characters of C1'-aglycons, have an influence on the stabilization of the 2'-oxyanion. The comparison of pK_a for 2'-OH in 1-deoxy-D -ribofuranose (**1a**) (13.56) with those of nucleosides adenosine (**1b**), guanosine (**1d)**, uridine (**1e**), cytidine **(1f**) (12.15-12.71) and 1-deoxy-D -ribofuranose 3',5'-O-bis-ethyl phosphate (**3a**) (14.05) with those of nucleosides adenosine 3',5'-bis-ethyl phosphate (**3b)**, guanosine 3',5'-methyl-ethyl phosphate (**3d)**, guanosine 3',5'- bis-ethyl phosphate (**3c)**, cytidine 3',5'-methyl-ethyl phosphate **(3f)** (12.99-13.53) clearly shows that the 2'-OH in adenosine (**1b**) and its 3',5'-bis-ethylphosphate derivatives (**3b**) is the most acidic (Table1). Similarly comparison of pK_a for 2'-OH in 3'-AMP (**2a**), 3'-GMP (**2b**), 3'-CMP (**2c**) and 3'-UMP (**2d**) (13.81 – 13.98) shows pK_a of 2'-OH in 3'-AMP to be lowest. This is because adenin-9-yl is a better stabilizer for the 2'-oxyanion compared to any other aglycons in the nucleoside, 3'-phosphomonoester, 3'-phosphodiester, or 3',5'-bis-phosphodiester series. From this work it has been established that the pK_a's for 2'-OH of the ribonucleosides and their phosphate derivatives simply change because of different abilities of various sugar substituents to stabilize the 2'-oxyanion.

Table 1.The pK_a values with standard error at 25°C for sugar 2'-hydroxyl dissociation in various nucleosides, their 3'-monophosphates, nucleotides, their abasic counterparts, as well as 3'-deoxyadenosine, *ara*-adenosine and aristeromycin. The pK_a is measured by the effect of pH on the ^{1}H chemical shift of the proton (H1', H2' and H3') for these compounds.

Compound name[#]	pK_a from H1'		pK_a from H2'		pK_a from H3'		Overall Average pK_a
	Titration curves Eq.1[b]	Calculated From Eq.2[c]	Titration curves Eq.1[b]	Calculated From Eq.2[c]	Titration curves Eq.1[b]	Calculated From Eq.2[c]	
1-d-rf (**1a**)	13.65±0.18	13.56±0.02	13.51±0.16	13.47±0.02	13.63±0.16	13.55±0.01	13.56±0.03
EtprfpEt (**3a**)[a]	14.04±0.12	–	14.15±0.14	–	13.96±0.12	–	14.05±0.06
Ado (**1b**)	12.10±0.04	12.07±0.03	12.22±0.06	12.10±0.03	12.23±0.05	12.15±0.04	12.15±0.03
ApEt (**4b**)	12.90±0.09	12.86±0.04	13.28±0.06	13.11±0.05	12.98±0.10	12.90±0.04	13.00±0.07
EtpApEt (**3b**)	13.07±0.08	12.87±0.04	13.24±0.09	12.95±0.04	13.00±0.08	12.83±0.04	12.99±0.06
3'-AMP (**2a**)	13.66±0.05	13.92±0.01	13.76±0.10	13.99±0.01	13.69±0.06	13.89±0.01	13.81±0.05
ara-Ado (**1g**)	–	–	12.80±0.16	12.83±0.10	12.58±0.16	12.61±0.08	12.71±0.06
3'-dAdo (**1c**)	12.90±0.06	12.59±0.10	13.23±0.11	12.98±0.06	13.36±0.13	13.07±0.04	13.02±0.11
Aristeromycin (**1h**)	13.37±0.11	13.13±0.02	13.49±0.13	13.18±0.03	13.34±0.10	13.14±0.02	13.28±0.06
Guo (**1d**)	12.53±0.03	12.56±0.03	12.63±0.05	12.76±0.05	12.45±0.03	12.61±0.02	12.59±0.04
MepGpEt (**3d**)	13.41±0.06	13.30±0.02	13.49±0.13	13.03±0.12	13.24±0.02	13.39±0.01	13.31±0.07
EtpGpEt (**3c**)	13.44±0.06	13.17±0.06	–	–	13.39±0.06	13.29±0.02	13.32±0.06
3'-GMP (**2b**)[a]	13.86±0.08	–	13.80±0.07	–	13.83±0.08	–	13.83±0.02
Urd (**1e**)	12.79±0.09	12.54±0.08	12.88±0.08	12.42±0.13	13.07±0.10	12.56±0.07	12.71±0.10
3'-UMP (**2d**)[a]	13.88±0.06	–	14.01±0.12	–	13.98±0.10	–	13.96±0.04
Cyd (**1f**)	12.45±0.05	12.45±0.03	12.53±0.03	12.55±0.02	12.60±0.05	12.55±0.02	12.52±0.02
MepCpEt (**3f**)	13.46±0.04	13.52±0.01	13.47±0.06	13.68±0.07	13.46±0.04	13.57±0.01	13.53±0.04
3'-CMP (**2c**)[a]	13.83±0.09	–	14.07±0.02	–	14.04±0.02	–	13.98±0.08

[a]*The pKa values of* **3a,2b, 2d** *and* **2c** *are estimated to be not less than 14.05, 13.83, 13.96 and 13.98 respectively.* [b] $pH = pK_a + \log[A^-]/[AH] = pK_a + (1-a)/a$.....*Eqn(1).*[c] $pK_a = pH + \log(\delta_h - \delta_{obs})/(\delta_{obs} - \delta_l)$.....*Eqn(2).* [#] *See scheme 1 for chemical formulae.*

2.3.3 The effect of aglycone on the pK_a of the internucleotidic 2'-OH group in diribonucleoside (3'→5') monophosphates

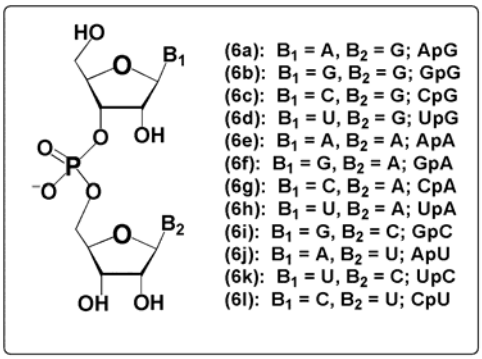
pH dependant pseudo first order rate constant (k_{mean}) for alkali mediated cleavage of diribonucleoside (3'→5') monophosphates under identical condition was used to determine the pK_a values for the internucleotidic 2'-OH of eight different diribonucleoside (3'→5') monophosphates (**6a-h,** Scheme 2). The diribonucleoside (3'→5') monophosphates (5'-N^1pN^2-3') were chosen such that while the 5'-N^1p (N^1 = A, G, C and U) was varied, the pN^2-3' have been kept constant, either to G or A, to examine if the nature of base-base stacking had any influence on the pK_a value of internucleotidic 2'-OH. It was seen that the pK_a values remain almost the same for all internucleotidic 2'-OH

Table 2. The pK_a values of diribonucleoside (3'→5') monophosphates and nucleoside 3'-ethyl phosphatse measured at 298 K and ionic strength of 1M NaCl

Compound	pK_a determination by pH dependant first order rate constant	pK_a determination by pH dependant ^{1}H chemical shift
ApG (**6a**)	12.71 ± 0.02	–
GpG (**6b**)	13.13 ± 0.04	–
CpG (**6c**)	13.17 ± 0.03	–
UpG (**6d**)	13.16 ± 0.03	–
ApA (**6e**)	12.81 ± 0.04	–
GpA (**6f**)	13.11 ± 0.08	–
CpA (**6g**)	13.28 ± 0.02	–
UpA (**6h**)	13.10 ± 0.03	–
GpEt (**4a**)	–	13.14 ± 0.06
ApEt (**4b**)	–	12.91 ± 0.05
CpEt (**4c**)	–	13.21 ± 0.04
UpEt (**4d**)	–	13.25 ± 0.03

for 5'-N^1p (N^1 = G, C and U) in dimers **6a-h**, except for the internucleotidic 2'-OH with 9-adeninyl as the 5'-aglycon [12.71 ±0.02 in ApG (**6a**) and 12.81 ±0.04 in ApA (**6e**)] which are 0.3 to 0.4 pK_a units more acidic compared to others shown in (Table 2).

Similar effect of nucleobases on the respective 2'-OH pK_a values were observed in monomeric nucleoside 3'-ethyl phosphates (**4a-d**) which are the simplest model of dimers with absence of any intramolecular base-base stacking. The pK_a of their 2'-OH were noted to be very comparable to those of the



Scheme 2

dimeric counterparts (Table 2) with the exception of ApEt (**4b**), in which pK_a of 2'-OH increases by 0.2 pK_a units compared to the pK_a of the internucleotidic 2'-OH in ApG (**6a**) and ApA (**6e**). It is to be further noted that in both the case of dimers (**6a** and **6e**) the internucleotidic 2'-OH group with adenosine as the aglycone at the 5'-end and the monomer (**4a**) having the 2'-OH groups with adenosine as the nucleobase showed the most acidic pK_a.

The internucleotidic 2'-OH of 5'-**A**p moiety of **A**pA (**6e**) or **A**pG (**6a**) is more acidic compared to the 5'-**G**p counterpart in **G**pA (**6f**) or **G**pG (**6b**) respectively. This is due to unique aromatic characters of their respective 9-adeninyl and 9-guaninyl groups. This is because the imidazole moiety in adenin-9-yl can fully donate its π charge successfully to the electron deficient fused aromatic pyrimidine moiety, which is only partly possible in the pseudoaromatic pyrimidine part in guanin-9-yl. Thus the imidazolyl moiety in guanin-9-yl system retains a considerable basic character by poorly conjugating its π charge to the fused 2-amino-6-pyrimidone moiety. The differential π-charge donating capacity of 9-adeninyl vis-à-vis 9-guaninyl allows the former to stabilize its 2'-oxyanion more efficiently compared to that of the latter, thereby causing an increased acidity of its 2'-OH group.

2.4 Implications

The study suggests that the pK_a of 2'-OH of different ribonucleotide units in a large RNA molecule can vary according to the local microenvironment and the hydrophobic character of the nucleobase, hence imparting different reactivities to the phosphodiester functions (sections 2.3.2 and 2.3.3). The differential stabilization of the internucleotidic 2'-oxyanion in nucleos(t)ides can also lead to differential hydration of 2'-OH. The fact that the absence of the ring oxygen in aristeromycin (**1h**) elevates the pKa of 2'-OH by 1.13 units compared to that in adenosine (**1b**) means that the substitution of pentose sugar by a cyclopentane moiety in an RNA molecule will make the general

acid-base-catalyzed 2'-OH assisted transesterification reactions slower in the former compared to the latter. Thus the importance of the pentoses prevails over the cyclopentane-based RNA-world.

3. Reactivity of phosphodiester group in RNA

After the discovery of RNA catalysis the hydrolysis/transesterification of RNA has been studied extensively in both model compounds[146-154] as well as in ribozymes[81-90,66,67,93]. It is still a challenge for the chemist to understand how RNA with a very few reactive functional groups compared to proteins can act as biocatalysts. To shed light into the mechanistic aspects of ribozyme catalysis, an endeavor was made to understand the mechanism of nonenzymatic cleavage of RNA and the factors affecting the kinetics and thermodynamics of the process. This knowledge can be utilized in the rational design of artificial catalytic ribonucleic acids.

3.1 Factors affecting nonenzymatic base-promoted degradation of RNA

Several factors have so far emerged as prerequisites in the nonenzymatic base catalyzed hydrolysis of RNA phosphodiesters[133,142,146-154]. They are (i) the nucleophilicity of the 2'-OH group which depends on its pK_a[133,142], (ii) The electrophilicity of the reacting phosphate[147,148], (iii) the in-line conformation of the attacking 2'-oxyanion with the developing 5'-oxyanion[149], (iv) readiness with which 5'-oxyanion leaving group departs[147,148], (v) the intramolecular environment[147-162], i.e stacking, hydrogen-bonding and nucleobase composition around the transesterification site. Thus a well-stacked rigid structure would retard the base-promoted cleavage of a RNA phosphodiester compared to a disordered structure. Similarly the rate of phosphodiester bond cleavage is different depending on whether the sequence undergoing degradation is within the loop of a hairpin or in the stem[153,162]. Due to H-bonding, the helical stem is more rigid and in-line attack is hindered compared to the loop region that behaves as a single stranded RNA. (vi) The hydrogen bonding network around the scissile bond, attacking the 2'-OH and the neighboring bases may accelerate the transesterification reaction in the following ways[159]: (a) by accepting a proton from the 2'-OH, (b) by donating a proton to the negatively charged phosphodiester or (c) by donating a proton to the leaving 5'-oxyanion. Direct or water-mediated hydrogen bond between the nucleobase and the non-bridging oxygen of 5'-phosphodiester has been shown to make the latter a better electrophile[163].

3.2 Importance of electron density around phosphate in RNA

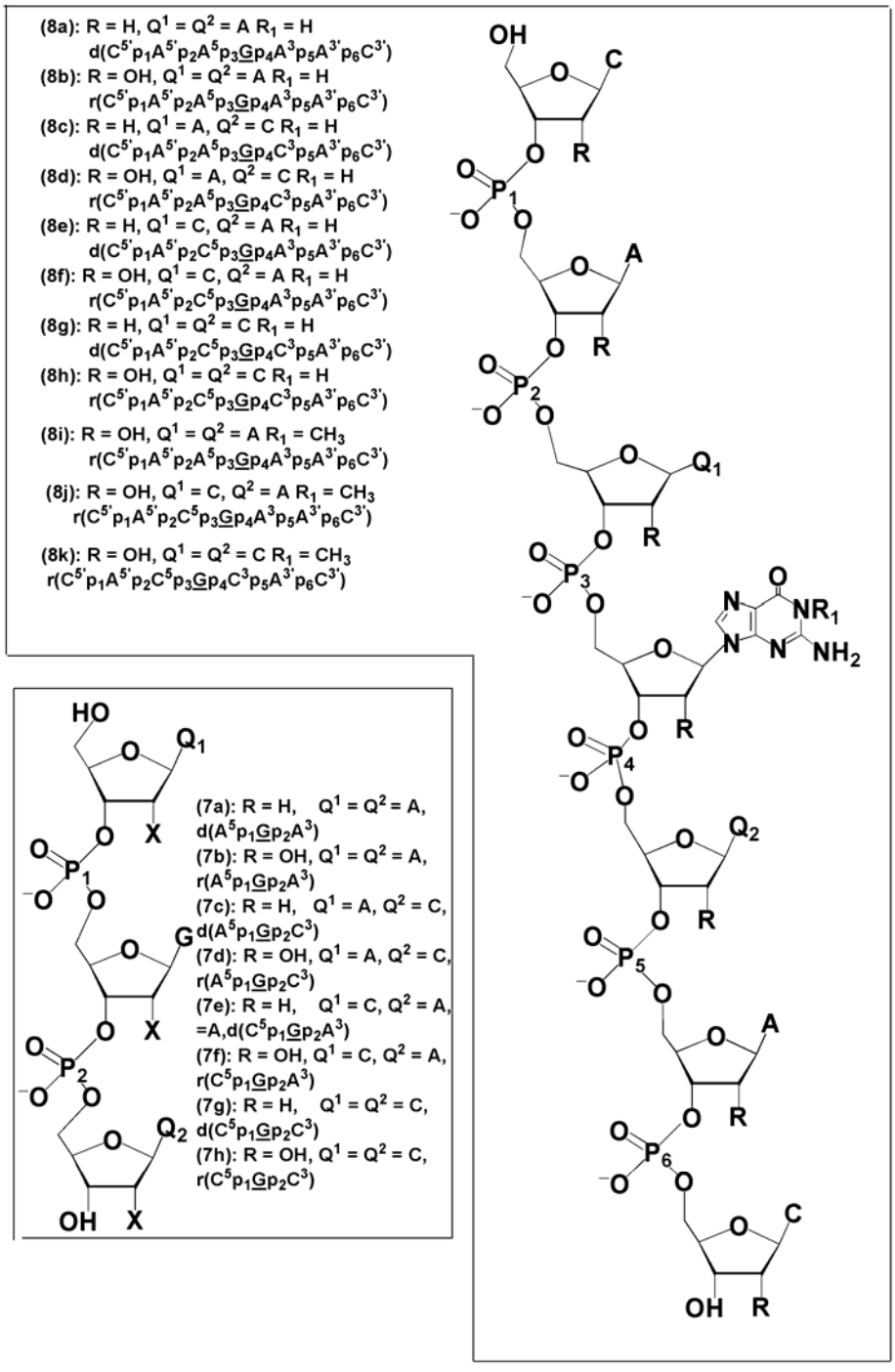
In ribozymes the enzymatic cleavage of phosphodiester bonds is achieved through precise substrate recognition, binding of the substrate to the enzyme by Watson-Crick base pairing[164-170,172], folding of the enzyme substrate complex using flexible domains (as in hammerhead) as well as tertiary interactions (as in Group I[164], Group II [165,166], HDV [167-169], Hairpin[168,169] and VS ribozymes [170]) to form the catalytic core. Following binding, the phosphodiester bond cleavage in ribozymes involves transesterification reaction (as discussed earlier in section 2.1.2) where the electrophilicity of the phosphate may play an important role in the attack of the scissile phosphate by a 2'-oxyanion. A similar importance of electrophilicity of phosphate may prevail in ribonuclease action on RNA that also proceeds through transesterification[67,92,93,171]. In the hammerhead ribozyme[172] it has been shown that a conformational change must occur at the phosphate center prior to cleavage to facilitate the in line attack of the nucleophile. In VS ribozyme [170] substrate two phosphate groups can act as ligand for two metal ions for metal dependant docking of substrate on ribozyme.

The electronic environment around the phosphate group in RNA is also vital for the interaction of proteins with RNA[173-176]. The crystal structure of glutaminyl tRNA synthetase (GlnRS)-$tRNA^{Gln}$ [173] complex show that hydrogen bonding between amino group of a guanine and a 5'-phosphate of an adenine in tRNA contributes to the recognition of GlnRS by $tRNA^{Gln}$. The RNA binding domain in sex-lethal protein[174] recognizes and binds uridine rich sequences by forming hydrogen bonds between 2'-OH groups and phosphate within the RNA.

3.3 Present work (Paper III)

The present work showed that the electronic environment around phosphodiester groups in ssRNA is non-equivalent in sequence specific manner whereas in ssDNA the phosphate group environments are rather comparable. This has been demonstrated by the pH dependant chemical shift (δ^1H as well as δ^{31}P) experiments of iso-sequential ssDNA/ssRNA followed by studies on the alkaline hydrolysis rates at specific phosphates in the ssRNA sequences.

3.3.1 Reflection of ionization of pseudoaromatic 9-guaninyl group on neighboring phosphate groups in ssDNA and ssRNA



Scheme 3

Titration at a particular pH range (6.6 - 12.5) on the model heptameric ssDNA and ssRNA sequences [d/r(5'-$Cp_1Ap_2Q^1p_3\underline{\mathbf{G}}p_4Q^2p_5Ap_6C$-3'): Q^1 = Q^2 = A (**8a**/**8b)** or C (**8g**/**8h**), Q^1 = A, Q^2 = C (**8c**/**8d**), Q^1 = C, Q^2 = A (**8e**/**8f**) Scheme 3] as well as trimeric ssDNA and ssRNA [d/r($Ap_1\underline{\mathbf{G}}p_2A$) (**7a**/**7b**), d/r($Ap_1\underline{\mathbf{G}}p_2C$) (**7c**/**7d**), d/r($Cp_1\underline{\mathbf{G}}p_2A$) (**7e**/**7f**), d/r($Cp_1\underline{\mathbf{G}}p_2C$) (**7g**/**7h**) Scheme 3] was used to generate a single negative charge at 9-guaninyl in the molecule in case of ssDNAs and trimeric ssRNAs. In case of heptameric ssRNAs, the formation of 9-guaninyl was accompanied by the simultaneous ionization of 2'-OH at pH values above 11.6. The effects of $\underline{\mathbf{G}}$/$\mathbf{G}^-$ and 2'-OH/2'-O^- ionization on the neighboring phosphate anions were studied by monitoring the change in chemical shift of ^{31}P markers of neighboring phosphate groups with changing pH. The anionic phosphates present in the oligonucleotide chain electrostatically interacted with the ionizing 9-guaninyl and 2'-OH group and as a result the ^{31}P markers showed the pK_a of guanine from the pH dependant ^{31}P chemical shifts.

3.3.2 Deshielding of phosphorus resonances in the alkaline pH

The internucleotidic phosphates in ssDNAs and ssRNAs are fully ionized at the studied pH range of 6.6 – 12.5 [69,70]. Hence the observed downfield shift of all ^{31}P resonances from the neutral to alkaline pH is a result of through-space repulsive electrostatic interaction of the phosphate anion and the $\mathbf{G}^-$ in trimeric ssDNA/ssRNA and heptameric ssDNA as well as the phosphate anion and the $\mathbf{G}^-$ /2'-O^- in heptameric ssRNA. The observed downfield shift of the ^{31}P resonances reflects weaker screening of ^{31}P nucleus owing to delocalization of charge into its dπ orbitals as $\underline{\mathbf{G}}$ becomes $\mathbf{G}^-$ over the pH range of 6.6 – 12.5 and 2'-OH becomes 2'-O^- at pH values 11.6 – 12.5 in heptameric ssRNAs (**8b**, **8d**, **8f** and **8h,** Scheme 3). This is very similar to the earlier observed downfield ^{31}P shifts in various types of phosphates[177-187], phosphonates[178,179] and aminophosphonates[179], as they are ionized with increase of pH.

3.3.3 Non-identical electronic environment around internucleotidic phosphates in ssRNA compared to isosequential ssDNA

In our work we define pK_{a1} and pK_{a2} as the pK_a value obtained by monitoring the change in δH8G and δ^{31}P respectively with pH for compounds **7a-h** and **8a-h**. In case of trimeric ssDNA (**7a**, **7c**, **7e** and **7g,** Scheme 3) and ssRNA (**7b**, **7d**, **7f** and **7h,** Scheme 3) as well as heptameric ssDNA (**8a**, **8c**, **8e** and **8g** Scheme 3) there is negligible difference between the pK_{a2} values (see Table 1 and Table S1 in **Paper III**) obtained from the marker phosphates and the pK_{a1} value obtained from δH8G of 9-guaninyl [the maximum difference

between the pK_{a1} and any of the pK_{a2}s within a particular sequence is 0.07 pK_a units in **7a** among trimers **7a-7h** and 0.14 pK_a units in **8e** among the ssDNA heptamers (**8a**, **8c**, **8e** and **8g**) which is very close to the error limit of ±0.13 pK_a units]. The pK_{a2} values within a given sequence in trimeric ssDNA and ssRNA as well as heptameric ssDNA are also not appreciably different [the maximum difference between the pK_{a2} obtained from the ^{31}P markers p_1 and p_2 is 0.02 pK_a units in **7d** and **7e** among ssDNA/ssRNA trimers (**7a-7h**). Among heptameric ssDNA (**8a**, **8c**, **8e** and **8g**) the maximum difference is 0.10 pK_a units between the pK_{a2}s obtained from the ^{31}P markers, p_2 and p_3 in **8g**]. Within the heptameric ssRNA sequences (**8b**, **8d**, **8f** and **8h,** Scheme 3) the pK_{a1} and pK_{a2} values significantly differ from each other in a sequence [the maximum difference between pK_{a1} and pK_{a2} (from ^{31}P marker p_2) is 0.84 pK_a units in **8h** and the minimum difference between the pK_{a1} and pK_{a2} (from ^{31}P marker p_4) is 0.05 pK_a units in **8b**]. In two ssRNA sequences, **8d** and **8f**, unfortunately we do not have the pK_{a1} of guanine from δH8 (**G**) to compare with the pK_{a2} obtained from ^{31}P marker. The pK_{a2} values obtained from each marker phosphates (p_2 –p_4) in a particular heptameric ssRNAs (**8b**, **8d**, **8f** and **8h**) also differ from each other depending upon the sequence context [among heptameric ssRNAs (**8b**, **8d**, **8f**, **8h**), the maximum difference between the pK_{a2} values obtained from the ^{31}P markers p_3 and p_4 in **8f** is 0.76 pK_a units and minimum difference is 0.25 p*K*a units between the pK_{a2} values obtained from the ^{31}P markers p_3 and p_4 in **8d**].

^{31}P markers of the internucleotidic phosphates proximal to the 3' and 5' ends of the ionization site $\mathbf{G}^-$ feel the effect of the ionization in the heptameric ssDNA and ssRNA. It has also been observed that only the phosphate markers p_2, p_3 and p_4 show the apparent pK_{a2} of guanine (see Figure 3, in **Paper III**).

3.3.4 Variable 9-guaninyl pK_a values from different phosphate markers in heptameric ssRNA

It has been observed that the pK_a of **G** obtained from the pH dependant ^{31}P chemical shift (pK_{a2}) of the internucleotidic phosphates in ssDNA/ssRNA (**7a-h**, **8a - h**) trimers and heptamers vary depending on the sequences and are different from the pK_a obtained from pH dependant δH8G chemical shifts (pK_{a1}) of the 9-guaninyl group in the trimers and heptamers itself. The reason for variation of the pK_{a2} can be explained as a result of variable electrostatic potential energy depending on variable phosphate charges and their distance from the charge generation site $\mathbf{G}^-$. The electrostatic potential energy $E = Q_1 * Q_2 / 4\pi\varepsilon_0 r$, where $Q_1 = \mathbf{G}^-$ and $Q_2 = PO_2^{-1}$; ε_0 = permittivity factor (depending on the microenvironment around each phosphates) and r = distance between charge generation site and phosphate. The relation suggests that as the distance r between the phosphate marker and the charge generation site $\mathbf{G}^-$

increases, electrostatic potential energy decreases, which means that the pK_{a2} from a given phosphate marker should decrease with respect to pK_{a1} from 9-guaninyl, provided all the internucleotidyl phosphates experience identical microenvironment. From the above relation it can also be justified that the nucleobases that are generally in a more hydrophobic microenvironment (i.e lower dielectric $4\pi\varepsilon_0$) compared to the internucleotidic phosphates, are expected to have larger pK_{a1} compared to the pK_{a2} from the phosphates. Hence the observation of pK_{a1}> pK_{a2} is well expected. On the other hand when a particular phosphate shows larger pK_{a2} compared to the pK_{a1}, a charge rearrangements owing to the electrostatic modulation of their electronic character by a hydrophobic local microenvironment[188,189], which has lower dielectrics compared to that of the aqueous environment is suspected to have occured. Thus, those specific phosphates which are in a more hydrophobic pockets than the others are likely to show relatively higher pK_a than the average pK_a normally measured for the internucleotide phosphates (1.5 to 2.1)[69,70] and will be relatively less ionized. Hence they are likely to be more electrophilic and will be easily attacked by nucleophiles such as 2'-oxyanion or the hydroxide. These phosphates can undergo transesterification reaction more readily in presence of 2'-oxyanion or the hydroxide ion.

3.3.5 Study of alkaline hydrolysis of heptameric ssRNAs in comparison with their **G** N^1-methylated counterparts

In order to investigate the relation between chemical reactivity at the internucleotidic phosphodiester bonds in heptameric ssRNA sequences and the modulation of electrostatic character at each phosphate centers due to formation of charged centers such as $\mathbf{G}^-$ and $2'O^-$ at alkaline pH values; alkaline hydrolytic cleavage at pH 12.5 and 20ºC was carried out for native heptameric ssRNAs (**8b**, **8d**, **8f** and **8h,** Scheme 3) and compared with the cleavage of the analogous N^1-**G**-methylated ($N^{1\text{-Me}}$-**G**) heptameric analogues (**8i**, **8j** and **8k,** Scheme 3) under an identical condition. In earlier instances[153-163] chimeric sequences with one reactive phosphodiester bond in single stranded and hairpin loops have been studied to show that the rate of non-enzymatic phosphodiester bond hydrolysis is sequence specific in nature. The hydrolysis of hepameric sequences **8b**, **8d**, **8f**, **8h, 8i**, **8j** and **8k** where more than one reactive phosphodiester is present at a time brought into light the inter-play of all the possible electrostatic interactions present in a real RNA sequence at an alkaline pH with ionized phosphate and 2'-OH groups. It has been found that (i) In case of the $N^{1\text{-Me}}$-**G** heptameric ssRNAs (**8i**, **8j** and **8k**), the total alkaline degradation occurs at a slower rate compared to that of the corresponding native ssRNA sequences (**8b**, **8f** and **8h**). (ii) In the three native heptameric ssRNAs (**8b**, **8f** and **8h** Scheme 3), alkaline hydrolysis is preferred to give the initial products at those internucleotide phosphates (p_2, p_3 and p_4, Figure 3 in

Paper III) which show both pK_{a2} and weaker screening of ^{31}P nucleus in alkaline pH compared to the neutral. These preferential cleavages found at the internucleotidic phosphates, p_2, p_3 and p_4 (which also show pK_{a2}) in the native heptameric ssRNAs, are 16-78% reduced in case of $N^{1\text{-Me}}$-**G** containing RNAs because of disappearance of the electrostatic effect of $\mathbf{G}^-$ (Figure 4 in **Paper III**). (iii) The percentile hydrolytic cleavage values at the internucleotidic p_2, p_3 and p_4 phosphates in **8k** is relatively less compared to those in **8h**, despite the fact the vicinal 2'-oxyanion population is considerably higher in the former (as evident from the $\Delta\delta^{31}$P shifts in Figure 5 **Paper III**). This suggests that the relatively high electrophilic character of phosphates in **8h** (contributed both by $\mathbf{G}^-$ and 2'-O^- in the proximity) is perhaps more important for its higher rate of the alkaline cleavage reaction than its 2'-oxyanion population (compared to that of **8h**), keeping in view that all other cleavage requirements in **8h** and **8k** is perhaps very similar because of closely similar sequence context.

3.4 Implications

Inequalities of the electronic environment around internucleotidic phosphates in RNA have tremendous implications in thorough understanding of the elementary reactions like the transesterification involved in non-enzymatic cleavage of RNA. This knowledge can be used in understanding the reaction mechanisms of protein nucleases as well. In small and large ribozymes on the other hand due to the difference in internucleotoidic phosphate charges caused by differential electronic environment (dictated by the sequence context, interaction with metal ion cofactor or by non-covalent interaction through distant neighboring group participation, folding pattern, varying hydration capabilities around each of the internucleotidic phosphates), a particular phosphate with elevated pK_a can have increased electrophilicity compared to other neighboring phosphates, which as a result may be more viable to transesterification reaction.

4 π-π interactions between stacked nucleobases in ssRNA

The various inter- and intra-molecular non-covalent forces, involving interactions of π systems within themselves as well as with other molecules, are considered important in chemical biology and material science. These forces are broadly classified as aromatic interactions[76,190-197]. The theoretical studies have shown the importance of these weak interactions involving in base-base stackings to stabilize the double helical structure of DNA[197]**.** Thses interactions are also present in drug-DNA interaction[197], DNA-protein interaction[197], in crystal packing of aromatic molecules[198], in stabilizing the tertiary structures of proteins[199], during binding of polyaromatic macrocycles[200-208], in many host-guest complexation[196,205] and also during the porphyrin aggregations[190,206].

4.1 Different types of aromatic interactions

The major non-covalent aromatic interactions, can be categorized as (i) π – π interactions,[76,207] (ii) CH – π interactions (involving C<u>H</u> of both aryl[202-204,208] and alkyl[209,210]) and (iii) ion – π interactions (involving both cation-π[211,212] as well as anion-π[213-216]). Various forces[190-196] stabilizing the aromatic stacking interactions include a) van der Waals interactions, b) electrostatics, c) dispersion d) charge transfer and e) solvophobic effect. According to Kool et al.[217] aromatic stacking interactions between nucleobases in water involve electrostatics (dipole-dipole and dipole-induced dipole) interactions, dispersion (momentary dipole-induced dipole) effects and solvation. The π-π aromatic interactions can also be categorized according to the geometries[76,190,194,196] of interactions as: edge-to-face (T-shaped), offset and face-to-face (Figure 5).

Although controversies exist regarding the most stable geometry of aromatic interaction, several studies[76,190-194] invoked that edge-to-face (Figure 5a) and offset stacked (Figure 5b) geometries are in general energetically favored over the face-to-face (Figure 5c) interaction. It was proposed by Hunter et al.[76,190,191] that during offset stacking and edge-to-face (T-shaped) interactions, attractive atom-πσ interaction (electrostatic interaction between π-electrons and σ-framework) was present rather than repulsive π-π interaction

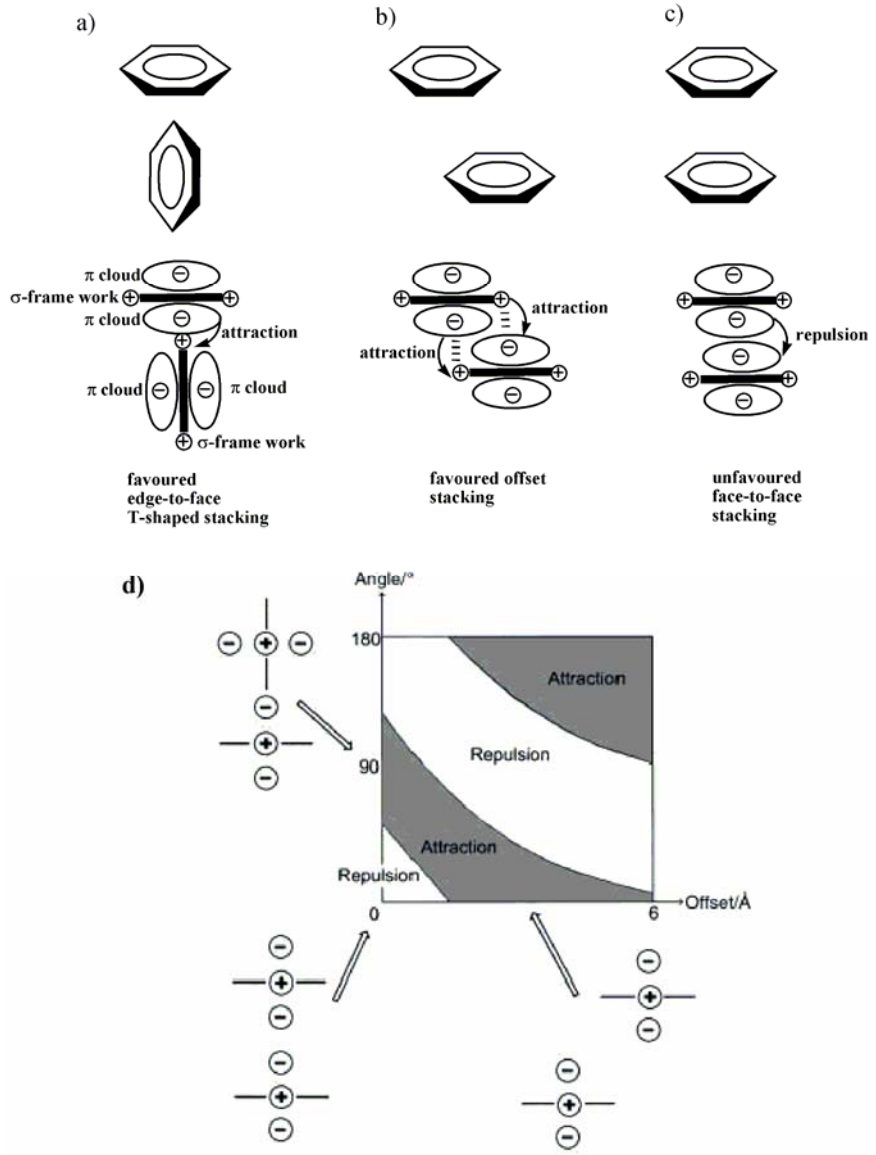


Figure 5. Panel a, b, and c show schematic representation of the geometry of aromatic interactions. The negative lobe represents the π-electron cloud and the σ-frame work is considered to be relatively positive. Panel d shows electrostatic interaction between π-charge distributions as a function of orientation of the aromatic moieties. The picture in panel d is adopted from ref 190.

in the face-to-face stacking. However, an attractive face-to-face aromatic interaction can be observed between donor-acceptor pairs of negatively charged π cloud of unsubstitued benzene ring and positively charged π framework of the hexafluoro benzene (C_6F_6)[207].

4.1.1 Predominating forces involved in inter- and intramolecular aromatic interactions

Temperature[202-204], concentration[201] and solvent[201,219] dependant studies on simple chemical model systems by NMR (intermolecular and intramolecular complexes) as well as computational methods[220] were employed to develop a thorough understanding of the different non-covalent forces (discussed in section 4.1) involved in the aromatic interactions and their extent of participation in the total interaction. Electrostatics has emerged as an important force in folding of model aromatic systems[76,190,191,211,212,222-228]or cation-π complexes[212,225]. Dispersion forces[220] are proved to be predominating in edge-to-face aromatic interaction in aryl CH-π, and relatively weaker alkyl CH

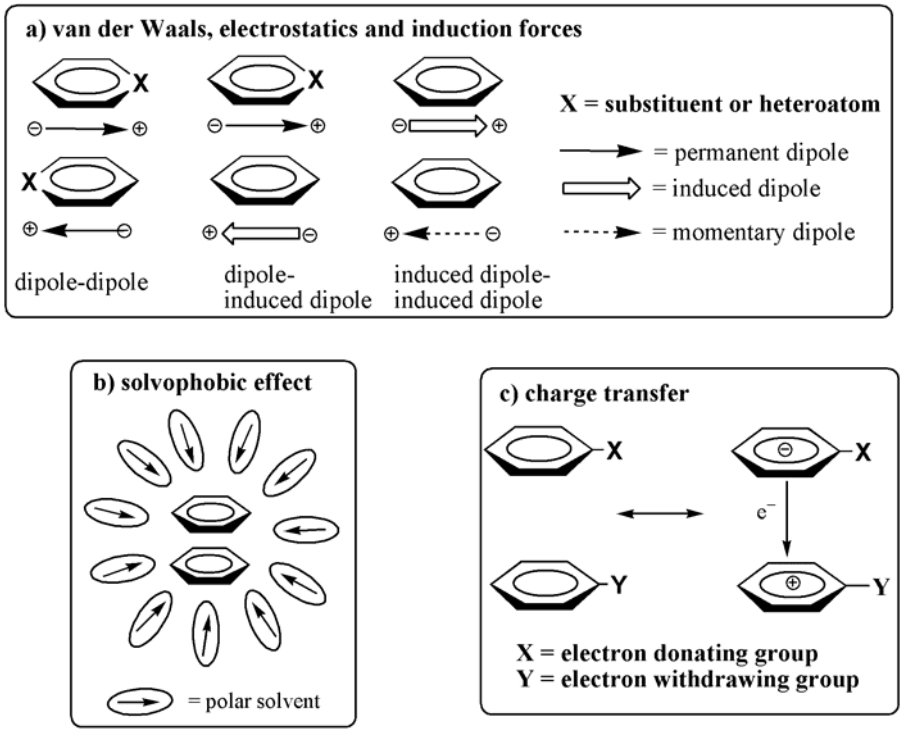


Figure 6. Panel (a) – (c) shows schematic illustration of different kind of non-covalent forces existing between two arene moieties.

-π interactions. Change in self association between offset stacked m-phenylene ethylene based cyclic and linear systems by changing the nature of substituents (electron donating or electron withdrawing) revealed that van der Waals interaction[218] have been found to predominate in these systems. Studies on the stacking interaction between representative hydrocarbon and heterocyclic aromatic moieties showed the predominance of polarizability of aromatic surface over hydrophobicity as major attractive components in such

interactions[201]. The study on foldamers[219] revealed that, desolvation of the aromatic surfaces provides the dominant driving force for complexation. Partial charge transfer has been proposed to arise from through-space proximity between alkyl hydrogen and aromatic moiety in hydrogen-bond like CH-π interaction[209] as well as in intra- and intertmolecular stacking[226] involving indole and adeninium rings. The dominance of polar/π interaction is observed in edge-to-face and offset stacked aromatic moieties in 1,8-di-o-tolylnaphthalenes[195,203,204] , in interactions of carboxylate ion and arene[229] as well as trimethylammonium ion and arene[230]

4.1.2 Aromatic interactions in nucleic acids

Aromatic interactions play an important role in pre-organizing the three dimensional structure in nucleic acids and their complexes[197,231-241] with protein. The presence of offset stacking between neighboring nucleobases in DNA[197] was observed by Hunter et al. from X-ray crystallographic data followed by computer modelling. Recent crystallographic and solution NMR studies on cytidine rich DNA fragments forming quadruplexes have proposed the presence of stabilizing O4' (lone pair) – π (nucleobase)[231] interactions. It has been found that H-π and lone pair-π interactions are present between water molecules and functionally important unstacked residues in the 1.25 Å crystal structure of the ribosomal RNA pseudoknot[232]. Cation-π interactions have been identified in protein-DNA complexes along with base stacking and hydrogen bonding in stabilizing stair shape motifs [233-235]. Cation-π interactions have been found between $Mg(OH_2)_6^{2+}$ and DNA bases in the major groove of distorted B-DNA[236]; the anticodon arm of yeast $tRNA^{phe}$ and to the magnesium core of the Tetrahymena group I intron P4-P6 domain. Cation-π interactions have been also implicated in DNA bending, DNA-protein recognition, base-flipping, RNA folding, and catalysis[236]. Recently thymine-methyl/π interactions have been shown to be important in sequence-dependant deformability of DNA[237]. The importance of electrostatic cation-π interactions was shown in the protein recognition of m^7G part of the mRNA cap by X-ray studies along with calorimetric and fluorescence analyses[238-240]. A key contact in the active site of aminoglycoside phosphotransferase enzyme has been shown to be a aromatic π–π stacking interaction between Tyr42 and the adenine ring of bound nucleotides[241].

4.2 Stabilization of nucleic acid structure by base stacking

It has been known for a long time now that nucleosides and nucleotides associate in polar solvent like water due to hydrophobic interactions with an association constant of $1M^{-1}$-$10M^{-1}$ [34]. As stacking interactions can play an important role in such associations, the thermodynamic parameters of stacking were studied in monomeric and oligomeric nucleic acids using vapour pressure osmometry,[34,251] temperature and pH-dependant NMR,[242-246,252] ORD[248] as well as theoretical simulations[247,249]. From these studies it was established that purine bases are in general more stacked than pyrimidine bases due to larger surface area for the former.

Detailed studies on the effect of dangling end bases on the stability of RNA/RNA[35,36,39] or DNA/DNA[37,38] or duplexes showed they stabilize duplexes [free energy of stabilization ($\Delta G°$) of both the DNA and RNA helix lies between -0.5 kcal mol^{-1} to -1.8 kcal mol^{-1}][253] through stacking interactions even in the absence of H-bonding. Duplex melting studies by NMR[35]and UV[36,38,39] brought into light various features of dangling end stabilizations. The above discussed inter- and intramolecular stacking interactions between nucleobases in monomeric and oligomeric nucleic acids have however failed to show any insight into the nature of their stacking interaction.

4.3 pK_a perturbation in biomolecules

The pK_a perturbation has important structural and functional[188,189,254-274] implications for biologically active nucleic acids and proteins. The pK_a perturbation can be a result of (a) interaction between charged groups to stabilize the structure (b) differential solvent accessibility of charged side chains and (c) H-bonding and stacking interactions between the functional groups to stabilize an enzyme substrate complex. The strengthening of H-bonding in the transition state compared to the ground state can be well achieved by decreasing the pK_a difference (ΔpK_a) between the H-bond donor and acceptor by being in a more hydrophobic environment (less dielectrics) in the transition state through charge rearrangement compared to the ground state[188,189]. Thus in general the shift in pK_a of a functional group from its original values in any system can provide information about neighboring charges and their electrostatic interaction, structural perturbation, and differential hydration of the microenvironment.

In contrast to proteins where histidine (pK_a = 6.0) can be found at the active site of enzymes quite often, nucleic acids do not contain functional groups with p*K*a's near physiological pH. Thus shift of pK_a values of RNA functional groups to near physiological pH (as has been observed in many ribozymes with adenine and cytidine nucleobase)[126-129,255-259,262] can diversify

RNA's function by making it a valid candidate for performing general acid-base catalysis. Biophysical and biochemical studies on leadzymes[256,257], DNA[260,261,263,264], proteins[265-273] and molecular receptors[274] revealed that fuctional groups have altered p*K*a's compared to their p*K*a's in the monomeric units according to the requirement of forming activated complexes for reactions. In some proteins[268-273] pK_a perturbation on the other hand plays a role in catalysis and structural stabilization.

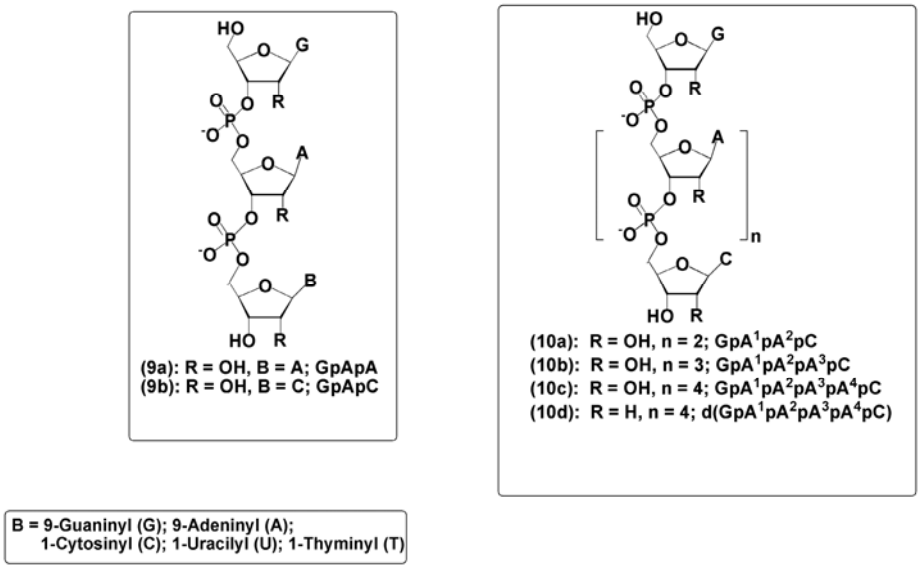
4.4 Present work (Papers IV - VI)

4.4.1 Cross-modulation between nucleobases of dimeric and oligomeric ssRNA

In absence of H-bonding stacking plays a more important role in the self-assembly process of ssRNA. It has been shown that the change in physico-chemical character of a nucleobase can be observed from its neighbor depending upon the orientation of adjacent stacked bases. This can have enormous implications in recognition of the single-stranded RNA which guide various fundamental life processes.

4.4.2 pH titration studies of dimeric and oligomeric ssRNAs

For the pH dependant chemical shift studies, a set of dimeric (**6a, 6f, 6h-6l** Scheme 2), trimeric (**9a, 9b** Scheme 4), tetrameric (**10a** Scheme 4), pentameric (**10b** Scheme 4) and hexameric (**10c** Scheme 4) ssRNA sequences were chosen such that only one among all the nucleobases present in the molecule could be exclusively protonated or deprotonated at a given pH range at 298 K. The pH dependant chemical shift change of the aromatic marker protons of the nucleobases in the dimeric and oligomeric ssRNA's showed that pK_a of a ionized nucleobase (protonated or deprotonated) could be obtained not only from its own maker protons, but also from the marker protons of the neighboring nucleobases. This observation was a result of cross-modulation of coupled π systems of the nearest-neighbor aglycones. The NMR studies on the ssRNA sequences (dimeric to hexameric) brought to light i) the nearest neighbor dependant change in pseudoaromatic character of the aglycones, ii) the extent of their propagation in a ssRNA sequence and iii) the various electrostatic interactions involved in the stacked nucleobases as well as iv) the thermodynamics of the offset stacking. The isomeric dimer pairs with all possible aglycone combinations *i.e* purine-purine (**6a** and **6f**), purine-pyrimidine (**6h** and **6j**) as well as pyrimidine-pyrimidine (**6k** and **6l**)



Scheme 4

showed the two way (3'→5' and 5'→3') cross-modulation of pseudoaromatic character of the nucleobase from the one to the other.

Chemical shift variation for an aromatic marker proton over the total pH range ($\Delta\delta_{N\text{-}D}$) was found to be a direct evidence of interplay of electrostatic forces affecting stacking ⇄ destacking among neighboring nucleobases. The upfield/downfield shift of the marker protons of the various nucleotide residues of oligo RNA's with pH ($\Delta\delta_{N\text{-}D}$) was due to more relative stacking/destacking of the nucleobases with the neighbors as a result of ionization of a single nucleobase in the oligomers (see Figure 3 in **Paper VI** for oligomers **6f, 6i, 9a, 9b, 10a-c**). Negligible change in chemical shift with pH for any marker proton of a nucleobase implied that some of those edges or parts of the constituent nucleobase is/are relatively destacked or bulged out throughout the pH range studied and hence weak electrostatic interactions exist between electron densities of these atoms with nearest neighbor nucleobases and phosphates. $\Delta\delta_{N\text{-}D}$ in compounds (**6f**, **6i**, **9a**, **9b**, **10a-c**) also showed that the repulsive electrostatic forces anion ($\mathbf{G}^{-}$)-π/dipole($Im^{\delta-}$) propagated from the first to the third nucleobase quite strongly in the ssRNA's, causing destacking of the helix. The repulsive force however was opposed by electrostatic attractive atom-πσ(major) and anion(G^{-})-π/dipole($Py^{\delta+}$) interactions. In ssRNA sequences it has been observed that δH8 (marker for the imidazole part of the 9-adeninyl) suffers larger deshielding upon deprotonation because of destacking owing to repulsive anion(G^{-})-π/dipole($Im^{\delta-}$) interactions. δH2 (marker for the pyrimidine part of the 9-adeninyl) shows only shielding in the deprotonated state because of attractive anion(G^{-})-π/dipole($Py^{\delta+}$) interactions.

4.4.3 Nearest-neighbor interaction between nucleobases and free energy of offset stacking in ssRNA

The partial charge on the nucleobase is altered on protonation or deprotonation as the pH of the medium is changed. The free energy change involved in the process corresponds to the pK_a of the nucleobase (ΔG_{pKa}) and is transmitted to the nearest-neighbor, modulating the pseudoaromatic character of the neighbors. The change in electronic character of the nucleobase in turn produced a shift in pseudorotational equilibrium of the sugar moiety (ΔG^o) (consistent with studies on monomeric nucleosides and nucleotides from the same lab)[23,275]. It is known that in RNA dimer, north-north pseudorotomers produce maximum stacking between adjacent nucleobases[243]. Thus by calculating the free energy of pseudorotation we could calculate the pH-dependant free energy of stacking (-0.3 kJ mol^{-1} ≤ $\Delta G^o_{stacking}$ ≤ -3.6 kJ mol^{-1} for dimers **6a, 6f, 6h, 6j-l, Paper IV**). The thermodynamics of the pH-dependant offset stacking in dinucleotides (**6a, 6f, 6h, 6j-l**) are shown in Figure 4D in **Paper IV**. All the dimers promote destacking with ionization of nucleobase, except in CpU$^-$ where the formation of 1-uracilylate as a function of pH failed to promote any destacking ($\Delta\delta_{N\text{-}D} > 0$, and $[\Delta G^o_{stackling}]_{N\text{-}state}$: –2.3 kJ mol^{-1} compared to $[\Delta G^o_{stackling}]_{D\text{-}state}$: –1.9 kJ mol^{-1}). In general, the extent of stacking decreases (see Figure 4 in **Paper IV**) in the following order: 5'-purine-purine-3' ≅ 5'-purine-pyrimidine-3' > 5'-pyrimidine-pyrimidine-3' > 5'-pyrimidine-purine-3' at the N-state. Calculation of free energies of pseudorotational equilibrium of individual sugar units in neutral and alkaline pH in ssRNA oligomers (**6f, 6i, 9a, 9b, 10a-c**) showed more south type conformation in deprotonated oligo RNA's proving the destabilization of the stacked RNA helix at alkaline pH upon guanylate ion formation. The difference in free energies between neutral and alkaline pH of the oligomers also proved that destacking in ssRNA oligomer has an effect through up to the third nucleotide in the tetra, penta and hexamer (see Figure 4 and Figure 5 in **Paper VI**).

4.4.4 An explanatory model for pK_a perturbation in single stranded oligonucleotides

The intrinsic pK_a of a nucleobase (**G** in our case) is the one found in a nucleotidic monomer unit with a single ionizable group along with the electrostatic interference of the constituent 3' or/and 5'-phosphates (which is already ionized, pK_a 1.5)[69,70], but in the absence of any nearest-neighbor nucleobase. In single stranded oligonucleotide sequences discussed in the present work, **G** (at the 5'-end of the sequence in present section and in the middle of the sequence in section 5) is one of the nucleobases which on ionization undergoes electrostatic interactions with additional neighboring electronic groups such as other nucleobases, phosphates, and the pentose-sugar units. This electrostatic interaction causes the modulation/perturbation of the intrinsic pK_a of the

G in the single stranded oligomers which is termed as the apparent pK_{a1}. This can be quantified by comparison with that of the respective monomer unit, **G**pEt (**4a**) or Etp(d/r**G**)pEt (**3j/3c**). Ideally, such a comparison of the apparent pK_{a1} of the **G** is best done within the subset of similar length oligomers thereby ensuring that the electrostatic effect due to the number of phosphates and sugars remains the same within the respective group. The marker proton from the neighboring nucleobases (**A** or **C**), which are non-titrable in the pH range 6.6-12.5, also shows the pK_a of **G** because of the variable electrostatic interaction of **G**/**G**$^-$ with the neighboring electronic groups which is termed as apparent pK_{a2}. Three distinct cases of variation of apparent pK_{a2} (with respect to apparent pK_{a1}), within each group of oligomers have been observed in this work: (i) when apparent pK_{a2} > apparent pK_{a1}, the electrostatic energy input from the corresponding reporter nucleobase owing to its nearest-neighbor influence and/or its location in a more hydrophobic microenvironment is suggested (**Type 2** effect); (ii) when apparent pK_{a2} < apparent pK_{a1}, an electrostatic screening by the solvent is suggested; (iii) when apparent pK_{a2} = apparent pK_{a1}, it means there is no additional electrostatic energy input from the electronic character of the neighboring nucleobase (**Type 1** effect). Both the apparent pK_{a1} and apparent pK_{a2} are site specific and depend on the relative difference in the microenvironments between the ionization site (at **G** to **G**$^-$ in our case) and the position of the neighboring nucleobases in the sequence. Both are the result of the two-state protonation ⇄ deprotonation equilibrium. The occurrence of apparent pK_{a2} at a distal site is a result of the electrostatic relay of the apparent pK_{a1} from the guanine ionization site. The actual transmission of pK_a is however modulated by the microenvironment and the resulting electronic nature of the nucleobase at the distal site. If the microenvironment around any of the distal nucleobases in a sequence is different, then the specific local hydrophobic environment around that nucleobase will (in comparison with the aqueous environment) have altered charge density and, as a result, will show different pK_{a2}'s in a sequence-specific manner.

4.4.5 Shift in pK_a value of the 9-guanylate in a ssRNA sequence compared to its corresponding monomer unit

The intrinsic pK_a of 9-guaninyl from δH8G in **G**$^-$pEt (**4a**) is 9.25 while the apparent pK_{a1} of 5'-**G**p/p**G**-3' (from H8G marker proton) residue in oligo-ssRNAs (**6a**, **6c, 6f, 6i,** Scheme 2) (**9a, 9b, 10a-c** Scheme 4) vary as follows: 9.42±0.01 in Ap**G**$^-$ (**6a**); 9.45±0.02 in Cp**G**$^-$ (**6c**); 9.17±0.02 in **G**$^-$pA (**6f**); 9.56±0.01 in **G**$^-$pC (**6i**); 9.75±0.02 in **G**$^-$pApA (**9a**); 9.88±0.02 in **G**$^-$pApC (**9b**);9.76±0.01 in **G**$^-$pApApC (**10a**); 9.82±0.01 in **G**$^-$pApApApC (**10b**) and 9.76±0.01 in **G**$^-$pApApApApC (**10c**) (Table 1 in **Paper VI**). This is due to the creation of differential microenvironment for the nucleobase, as a result of modulation of pseudoaromaticity through nearest neighbor offset stacking

interaction when placed in different sequence context. The nucleobase (9-guaninyl) in trimers (**9a, 9b**), tetramer (**10a**), pentamer (**10b**) and hexamer (**10c**) shows more basic pK_a compared to monomer (**4a**) or dimer (**6a, 6c, 6f, 6i**) because of both the neighboring nucleobases and phosphates (Table 1 in **Paper VI**). The pair wise comparison in dimer and trimer level show: pK_a of 9-guaninyl in **G**$^{-}$pA (**6f**, 9.17 ±0.02)/**G**$^{-}$pC (**6i**, 9.56 ±0.01) giving $[\Delta pK_a]_{(6i)-(6f)} \approx 0.39$. Similarly **G**$^{-}$**p**ApA (**9a**, 9.75 ±0.02) and **G**$^{-}$**p**ApC (**9b**, 9.88 ±0.01) shows $[\Delta pK_a]_{(9b)-(9a)} \approx 0.13$. Thus, the p$K_a$ of N^1-H of 9-gunaniyl residues is sequence dependant since the two set of dimers and/or trimers have the same phosphate charge but different nucleobase at the 3' or 5' end of **G** and their pK_a differs. Moreover the trimers (**9a, 9b**) have the same phosphate charge but there is a slight difference in the pK_a of 9-guaninyl residue $[\Delta pK_a]_{(9b)-(9a)} \approx 0.13$ which can be attributed to the difference in nucleobase sequence. This also 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.

4.4.6 Different pK_a value in nucleobase due to dissimilar electrostatic interaction in 3'- versus 5'-phosphate

Comparison of the pK_a of 9-guaninyl in GpEt (**4a**) (9.25 ±0.02) and EtpG (**5a**) (9.57 ±0.01) shows that the 5'-phosphate makes the pK_a of 9-guaninyl more basic ($[\Delta pK_a]_{(5a)-(4a)} \approx 0.32$) compared to the 3'-phosphate (Table 1 in **Paper VI**). 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.

4.4.7 Variation in the pK_a values of 9-guaninyl as obtained from different marker protons of nucleobases across the ssRNA single strand

The variation of apparent pK_{a2}s (ΔpK_a) for 9-guaninyl ionization as measured from the aromatic marker protons of other nucleobases *i.e.* A's and C's (Table 1 in **Paper VI**) across the strand in ssRNA are: $\Delta pK_a = 0.85$ for hexameric (**10c**); $\Delta pK_a = 0.48$ for pentameric (**10b**); $\Delta pK_a = 0.15$ for tetrameric (**10a**) ssRNA respectively. Similar ΔpK_a's for trimeric ssRNA are relatively small. Thus, the corresponding $\Delta\Delta G^{o}_{pK_a}$ values (in kJ mol^{-1}) which are 4.8 for **10c**; 2.8 for **10b** and 0.9 for **10a** have been attributed to the variable strength of electrostatic interactions of the offset stacked nucleobases among themselves as well as with the phosphates, across the strand in ssRNA. The variable pK_a

values of 9-guaninyl ionization as obtained from marker protons (H2A, H8A, H5C, H6C) of other nucleobases show that the **Type 2** effect predominates in ssRNA oligomers.

4.5 Implications

The pH dependant studies on ssRNA's help to understand how the change in electronic environment brought about by protonation or deprotonation of a particular constituent nucleobase can affect microenvironment of its own and the neighboring nucleobases, thereby influencing the self-assembly process of the ssRNA in general. This study shows that the change of microenvironment around a certain nucleobase propagates across the nucleotydic strand in variable manner depending upon the nature of nearest-neighbor aromatic interactions. The anionic and cationic centers generated in the ssRNA on changing the pH mimic the change of the environment due to binding of a ligand to the ssRNA. The ligand binding involves week non-covalent forces like electrostatic, hydrophobic and van der Waals as well as relatively strong forces like H-bonding and stacking between interacting moieties. The modulation of the pseudo-aromatic character of the nucleobase caused by environment dependant changes can hence dictate the ligand binding properties by steering the strength of molecular interaction forces. The spread of electrostatic interactions due to ssRNA, and/or ligand interaction can be controlled by the neighboring nucleobases depending upon whether they are electronically coupled i.e (stacked or not) thereby acting as an on-off switch in such molecular recognition process.

5 Sequence specific recognition in nucleic acids

It has been found that both ssDNA and ssRNA sequences are involved in protein recognition through sequence specific binding in many biological process like transcription, replication and recombination. Besides this, DNA and RNA aptamers[63-65,289-301] also bind to a large variety of ligands (like co-factors, amino acids, nucleotides, saccharides, antibiotics, peptides, proteins and other nucleic acids) with high sequence specificity and affinity. The sequence specific interactions in nucleic acids can be classified as i) ssDNA-protein interactions[276-282] and RNA-protein interactions[283-288] and ii) Aptamer binding to ligands[63-65,289-301]. Thus studies on sequence specific interactions in nucleic acids guide various cellular processes as well as shed light on structure recognition and catalytic function of nucleic acids.

5.1 Sequence specific interactions of ssDNA and ssRNA with proteins

Many ssDNA have shown their functionalities on binding to specific proteins in a sequence specific manner[276-282]. The DNA binding proteins include i) gene regulatory proteins[276,277,280] ii) proteins for which DNA is a substrate[278,279] like polymerase, topoisomerase and recombinase. The single stranded regions (sequence containing 5'-GTACCACC-3') in promoters of several cellular and viral genes interact with high affinity towards transcription factors [276] and optimize rates of transcription. The ssDNA binding domain of the protein Cdc13[277], shows high sequence specificity for GT rich sequences in yeast telomeres. The RecA[278,279]protein in the presence of ATP binds preferentially to the less stacked polypyrimidine ssDNA sequences. On the other hand, the presence of stacked purine rich sequences obstructs recA polymerization. Autoantibody 11F8 sequence[282] specifically recognizes a 19mer ssDNA sequence which has TCC as the tenth, eleventh and twelvth residues respectively. The sequence specificity in DNA-protein interaction may arise from designated interaction of protein with a specific DNA structural variant which is characteristic of a specific sequence.

Several ssRNA[283-288] have also shown to interact sequence-specifically with proteins. The interactions between RNA and protein is through i) van

der Waals contacts ii) hydrogen bonding, iii) salt bridges and iv) stacking interactions. Examples of such interactions can be found in sex-lethal proteins[284] binding to uridine rich sequence (UGUUUUUUU) of one of the introns of mRNA precursor; in the recognition of 3'mRNA poly(A) tail by poly(A) binding protein (PABP)[285] ; in transcription termination factor Rho[287] protein in its associates with two consecutive cytidine nucleotides and in the tRNA binding attenuation protein (TRAP)[286] that can bind to target sequence within ssRNA.

5.2 Sequence specific ligand binding of aptamers

The ligands recognized by aptamers tend to be planar, contain H-bond donating and accepting groups as well as positively charged groups. Aptamers on the other hand generally use purine rich sequences engaged in non-canonical base pairing interactions and tend to arrange for proper surfaces and H bond donors and acceptors for proper ligand binding. Irregular chain topologies, across helix stacking along with H-bonding as well as hydrophobic and electrostatic interactions stabilize the active conformations of the complex. The RNA aptamers include sequences binding to nucleotide cofactors like FMN[63], AMP[63] and ATP[63]; amino acids like L-Arginine[63] and L-Citrulline[63], drugs like Theophylline[63], and antibiotics like Tobramycin[63] and Neomycin B[63]. DNA aptamers on the other hand binds to thrombin[64], HIV-I reverse transcriptase[64], ATP[63], human neutrophil elastase (HNE)[64] and arginine[63]. G-quartet[65] and i-motifs have been identified as an important motif involved in ligand binding in all of these DNA sequences. Single stranded DNA (DNAzymes)[292-300] have also been shown to form specific binding pockets similar to RNA and to catalyse chemical reactions. Aptamers are selected from large libraries of DNA and RNA by invitro selection procedure[289,290].

5.3 Present work (Paper VII)

Importance of sequence-specific interactions in nucleic acids is evident from our studies described in sections 5.1 and 5.2. Sequence specificity is an intrinsic property of single stranded nucleic acids and has been shown even in small heptameric ssRNA and ssDNA sequences. The chemical reactivity of single stranded nucleic acids is thus expected to be strongly modulated by sequence context.

5.3.1 pH titration studies on trimeric and heptameric ssRNA as well as ssDNA

Heptameric and trimeric ssDNA/ssRNA molecules (**8a-h** and **7a-h,** Scheme 3) are designed in a way that, among all the aromatic residues present, only a single nucleobase (central 9-guaninyl moiety) ionizes at alkaline pH. Variation of electronic environment around 9-guaninyl has been studied depending on the sequence context of oligonucleotides. Efficiency of the electrostatic propagation of ionization of 9-guaninyl ($\mathbf{G}^-$ formation) through the neighboring nucleobases has been investigated in various ssDNA or ssRNA sequences. The trimeric sequences (**7a-h** Scheme 3) were used for comparison with heptamers to show how the electrostatic effect of the $\underline{\mathbf{G}^-}$ in the trimeric sequences is modulated when it is inserted in to a larger oligomer with an altered sequence context.

5.3.2 Sequence dependant pK_a modulation of the central 9-guaninyl (pK_{a1}) in heptameric ssRNA and ssDNA sequence

The apparent pK_a (pK_{a1}) of ***G*** in ssDNA trimers (**7a**, **7c**, **7e**, **7g** Scheme 3) varies from 9.49 ± 0.01 to 10.24 ± 0.01 (ΔpK_{a1}= 0.75); in ssDNA heptamers (**8a**, **8c**, **8e**, **8g** Schene 3) varies from 10.39 ± 0.01 to 11.06 ± 0.01 (ΔpK_{a1}= 0.67); in trimeric ssRNAs (**7b**,**7d**,**7f**,**7h** Scheme 3) varies from 10.03 ± 0.01 to 10.25 ± 0.01 (ΔpK_{a1}= 0.22) whereas in heptameric ssRNAs **(8b**, **8d**, **8f**, **8h** Scheme 3**)** it varies from 10.09 ± 0.02 to 10.58 ± 0.01 (ΔpK_{a1}= 0.49). The variation in apparent pK_a (ΔpK_{a1}) of $\underline{\mathbf{G}}$ observed within the set of the trimeric or the heptameric ssDNAs/ssRNAs is a measure how variable its pseudoaromatic character is within each group as a result of variation of the respective sequence, while the phosphate charges or/and the pentose sugar units within each group of trimers or heptamers remains the same. When the apparent pK_a (ΔpK_{a1}) variation of the $\underline{\mathbf{G}}$ in ssDNA trimers (**7a**, **7c**, **7e**, **7g**) is compared with the corresponding ssDNA heptamers (**8a**, **8c**, **8e**, **8g**) and, the ssRNA trimers (**7b**, **7d**, **7f**, **7h**) with the corresponding ssRNA heptamers (**8b**, **8d**, **8f**, **8h**) the effect of different microenvironments owing to their respective nearest-neighbors depending upon the chain length, phosphate charges or/and the pentose sugar units, as well as the sequence-context as a result of variable stacking is observed.

The apparent pK_{a1} of $\underline{\mathbf{G}}$ in heptameric ssDNAs **(8a**, **8c**, **8e**, **8g)** and ssRNAs **(8b**, **8d**, **8f**, **8h)** is always more basic compared to the pK_a of $\underline{\mathbf{G}}$ in their monomeric counterparts EtpdGpEt (**3j**) and EtpGpEt (**3c**) respectively (Table 1 **Paper VII**). By plotting such a difference in pK_a of $\underline{\mathbf{G}}$ [i.e., ΔpK_a, where ΔpK_a = {(pK_a of $\underline{\mathbf{G}}$) from δH8G in ssDNA or ssRNA} - {(pK_a of $\underline{\mathbf{G}}$) from δH8G in EtpdGpEt or EtpGpEt}] as a function of either (i) the chemical shift of $\underline{\mathbf{G}}$ (see Figure 3a in **Paper VII**) or (ii) the oligomerization shift (see Figure 3b in **Paper VII**) (difference in chemical shift between the monomer

and oligomer), within the series of heptameric ssDNA and ssRNA sequences (in which the number of phosphate charges remain the same) at the neutral pH, we have shown the effect of the sequence context promoted modulation of the respective chemical environments. Both these plots show a correlation with a high correlation coefficient (0.9 and 0.85, see Figure 3 in **Paper VII**). The chemical shift (δH8) and the oligomerization shift of **G** are manifestations of the strength of stacking of **G** with its nearest neighbors as well as the sugar-phosphate backbone conformation around **G** in the heptameric ssDNAs and ssRNAs. *Such high correlation reveals that the pK_a perturbation (ΔpK_a) of* **G** *in any oligomer increases with its increasing stacking interaction with the nearest-neighbors.* The most up-field shifts at the N-state as well as higher oligomerization shifts for **G** occur in the sequences with 5'-purine(A)-**G**-purine(A)-3' in the middle (as in the heptamers **8a** and **8b**) compared to all ssDNA/ssRNA sequences; and the least up-field shifts at the N-state as well as lowest oligomerization shifts for ***G*** in sequence having 5'-pyrimidine(C)-***G***-pyrimidine(C)-3' in the middle (as in the heptamers **8g** and **8h**) compared to all ssDNA/ssRNA sequences show that the former sequence ***G*** is most stacked and in the latter sequence ***G*** is least stacked among the heptameric ssDNA and ssRNAs (**8a-h**). However, in the case of trimeric ssDNA and ssRNA sequences (**7a-h**), such correlations could not be obtained, as their structures are presumably more random than those of the heptamers.

5.3.3 The electrostatic cross-modulation and pK_a perturbation of 9-guaninyl (pK_{a2}) among neighboring nucleobases in heptameric ssRNA and ssDNA sequence

The variable apparent pK_{a2} obtained from marker protons of nucleobase moieties, **A** and **C** in trimers (**7a**, **7c**, **7e**, **7g,** Scheme 3**)** and (**7b**, **7d**, **7f**, **7h,** Scheme 3) and heptamers (**8a**, **8c**, **8e**, **8g,** Scheme 3) and (**8b**, **8d**, **8f**, **8h,** Scheme 3) is the result of cross-modulation of pseudoaromatic character of centrally located ***G*** residue to the neighbors, owing to tandem nearest-neighbor electrostatic interactions. This gives rise to **Type1** effect (pK_a of 9-guaninyl from the neighboring nucleobases is the same as 9-guanylate ion) as the most predominant type of electrostatic interactions among the coupled nucleobases in all trimeric ssDNAs and ssRNAs and heptameric ssDNA sequences. This is because the pseudoaromatic characters of **A** or **C** in the above sequences have been similarly modulated by the electrostatic interaction of the G^- (apparent pK_{a2}, **Type 1** effect) which means that the pseudoaromatic characters of **A**s and **C**s at both 3'- and 5'-ends are uniform. Both **Type 1** and **Type 2** effects (pK_a of 9-guaninyl from the neighboring nucleobases are larger than that of 9-guanylate ion, owing to their own intrinsic electronic property) were however observed in the apparent pK_{a2} for the heptameric ssRNAs (**8b**, **8d**, **8f**, **8h**) depending on the respective sequence con-

text (see Table 1 **Paper VII**). Heptameric ssRNA sequence r(5'CAA***G***AAC-3')(**8b**) (***G*** flanked by 5'-**A** and 3'-**A**) showed comparatively poorer (ΔpK_{a2} = 0.23) cross-modulation of pK_a of ***G*** by the neighboring nucleobases (**Type 1** effect) while the sequences r(5'-CAA***G***CAC-3') (**8d**) (ΔpK_{a2} = 0.56, ***G*** is flanked by 5'-**A** and 3'-**C**), r(5'-CAC***G***AAC-3') (**8f**) (ΔpK_{a2} = 0.43, ***G*** is flanked by 5'-**C** and 3'-**A**), and r(5'-CAC***G***CAC-3') (**8h**) (ΔpK_{a2} = 1.25, ***G*** is flanked by 5'-**C** and 3'-**C**) show higher crossmodulation of pK_a, owing to the variation of the apparent pK_{a2} (**Type 2** effect) as a result of sequence-dependant nearest neighbor effect. *Thus it is evident that the pK_{a2} perturbation within the sequence is maximum when the neighboring nucleobases of* ***G*** *are pyrimidines, whereas it is minimum when the neighboring nucleobases of* ***G*** *are purines.* Moreover the variation of the pK_a of ***G*** as measured from the marker protons of **A**s (apparent pK_{a2}, **Type 2** effect) shows the modulation of the pseudoaromatic character of **A**s in (**8d)**, (**8f)**, and (**8h) (**Table 1 **Paper VII).** Thus the 3 A's in (**8d**) are not identical, nor are the 3 A's in (**8f**) nor the 2A's in (**8h**)**.** *This is due to the different partial ionic charges for the adenines owing to the differences in their intrinsic pseudoaromatic characters depending upon their nonidentical microenvironments.* The pseudoaromatic characters of **C**s have also been assessed by comparing the pK_a's of ***G*** as measured from the neighboring marker protons of **C**s (apparent pK_{a2}, **Type 2** effect). *This also shows that the pseudoaromatic characters of 3C's in* (**8d**)**,** *3C's in* (**8f**) *and 4C's in* (**8h**) *are electronically non-uniform.*

By comparing the propagation of electrostatic interaction in ssDNA/ssRNA sequences, it has been found that for trimeric ssDNA (**7a**, **7c**, **7e**, **7g**) and ssRNA (**7b**, **7d**, **7f**, **7h**) sequences electrostatic transmission was observed to be similar both at the 3'- and 5'-direction of ***G***. In contradistinction, the transmission of electrostatics interaction (manifested in apparent pK_{a2} obtained from the neighbors) of ***G*** in 3'- and 5'-directions in heptameric ssDNAs/ssRNAs is different depending upon the sequence context in the deoxy versus ribo series. The distance up to which the effect of electrostatics of $\boldsymbol{G}^-$ is transmitted through the neighboring nucleobases (as evident by observation of apparent pK_{a2}) is shown in Figure 2 **Paper VII**.

5.4 Implications

The result of transmission of electrostatic properties of an ionized nucleobase to its neighbors as well as the interactions between the base-base stacked neighboring aglycones in a single stranded nucleic acid sequence is the creation of a *unique set of aglycones* in an oligo or polynucleotide. The physico-chemical properties of these nucleobases are completely dependant upon the nearest neighbor electrostatic interactions. Thus in a RNA sequence, $P^1Q^1\underline{\mathbf{N}}Q^2P^2$, 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 $\leftrightarrows$ 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 depends on the intrinsic dynamics of folding and unfolding within the molecule owing to the sequence context or interaction with an external ligand. This has considerable implication in the specific ligand binding ability, aptamer recognition, RNA catalysis, and most probably in codon-anticodon interaction. 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 pK_a variation), which is manifested in its dual biological role in general, as we witness in the translation machninery and catalysis.

6 Importance of single-stranded and duplex structures in functioning of nucleic acids

6.1 Role of single- and double-stranded nucleic acids in biological process

H-bonding, stacking and electrostatic interactions[28-39,63-65] are important features in nucleic acid structure and recognition. In the single stranded form of nucleic acids the active groups available for interactions are well exposed. Recognition mainly takes place through single stranded regions in RNA and DNA, however double stranded DNA has been known to be recognized by ligands through intercalation[302,303] and helical interactions in the minor[304] and major groove[305].

A number of examples regarding importance of single-stranded nucleic acid in RNA structure lies in the stabilization of the cloverleaf structure of tRNA[306] by the 3'-terminal CCA end of tRNA, which is involved in aminoacylation reaction. Codon-anticodon recognition takes place through base pairing between tRNA and mRNA single strands[307]. A dangling nucleotide[35-39] at the end of both DNA and RNA duplexes is known to increase the duplex stability (discussed in section 4.2). A dangling nucleotide at the 3'-end is known to stabilize one of the stems in the RNA pseudoknot structure[308]. Single unpaired base bulges in RNA duplexes enhance[309,310] the stability of the RNA more compared to the fully base paired counterpart. The catalytic activity of Group I introns depend on recognition of a short complementary sequence called the internal guide sequence (IGS) in the enzyme by the substrate by base pairing to form the P1 helix[164,165]. Recognition and interaction with many ligands including proteins also take place with ssRNA and ssDNA[276-280,282,311-319] (as discussed in section 5.1) The ssDNA binding proteins belong to either of the two following categories: (i) those that recognize a particular sequence of nucleic acids like the transcriptional regulation proteins,[276] and the proteins involved in telomere replication[277] and (ii) those that specifically interact with the minor and major grove of DNA[278-280] like *Escherichia coli* ssDNA binding protein, RepA protein[280] or RecA protein[278,279]. The stacking interaction between aromatic side chain of proteins and nucleobaes are considered to be a major force in the binding process[320] for proteins undergoing the second kind of interactions. Stacking between aromatic amino

acids and nucleic acid bases plays an important role in the specificity of the enzyme for its complex with nucleic acid substrate. [281]

The most abundant secondary structural motif of nucleic acids is the homo (DNA-DNA/RNA-RNA) and hetero (DNA-RNA/RNA-DNA) duplex formation. Duplexes are an integral part of cellular processes like transcription and DNA replication, as well as intermediates in the reverse transcription of retroviral RNAs like HIV. Duplex stability is also an important factor in the design of gene silencing agents[24,321-323] such as, trans-cleaving ribozymes[321], small interfering RNAs[321] and antintisense oligonucleotides[24,321-323] where RNA/RNA or DNA/RNA duplex formation is a necessary step for degradation of the targeted mRNA.Thus an in-depth understanding of the conformational preorganization of the ssDNA and ssRNA as well as the factors affecting stability of duplex structures is a pre-requisite for understanding the various intra and intermolecular functionalities.

6.2 Characterization of single-stranded nucleic acid structures using NMR spectroscopy and NMR constrained Molecular Dynamics

Three types of regular helical pattern are commonly found in solution for DNA/DNA and RNA/RNA duplexes: A-, B- and Z-type, as discussed previously in section 1. By analyzing the relative intensities of aromatic to sugar residue (n, n+1, n-1) inter- and intra-residue NOESY[324] cross-peak volumes (distance between two protons can be estimated from cross-peak volume) as well as $^{3}J_{1H\text{-}1H}$ and $^{3/4}J_{1H31P}$ coupling constants valuable information can be obtained regarding helical characteristics, backbone and sugar puckering in nucleic acid duplexes. A schematic representation of the intra and inter proton contacts less than 5Å between nucleotide units in one of the strands of duplex are shown in Figure 7a for A-, B- and Z-type DNA/DNA or RNA/RNA duplexes. Due to the presence of syn guanosaine in Z-type DNA/DNA or RNA/RNA duplex, the intensity of intra-residue cross-peak between H1' and H8G (2.2 Å) is much stronger than that in A-and B-type DNA/DNA or RNA/RNA duplex. The inter residue $(H8/H6)_n$–$(H1'/H2'/H3')_{n-1}$ cross-peaks on the other hand is stronger (2 – 5.1 Å) for A and B type duplexes compared to that of Z-type duplex ($(H8G)_{n+1}$-$(H1'/H2'/H3')_n$ is 6.9 to 8 Å). The inter residue cross-peak intensity $(H8)_{n+1}$ –$(H4'/H5'/H5'')_n$ is stronger (3.3 to 4.7 Å) in guanosine moieties of Z-type duplexes compared to A- and B-type duplexes (6.9 to 7.2 Å). A- and B-type DNA/DNA or RNA/RNA duplexes can be differentiated from each other on the basis of relative intensity comparison of intra residue $(H8/H6)_n$ – $H1'_n$ and $(H8/H6)_n$ – $H3'_n$ as well as inter residue $(H8/H6)_n$ – $H1'_{n-1}$ and $(H8/H6)_n$ – $H3'_{n-1}$ cross-peak. In A type helix as the sugar is in 3'-endo conformation the$(H8/H6)_n$ – $H3'_{n-1}$/ $(H8/H6)_n$ – $H3'_n$

cross-peak is stronger (~3Å) than$(H8/H6)_n - H1'_{n-1}$ / $(H8/H6)_n - H1'_n$ cross-peak. In B type helix the sugar is in 2'-endo conformation and hence $(H8/H6)_n - H1'_{n-1}$/ $(H8/H6)_n - H1'_n$ cross-peak is stronger (~3Å) than $(H8/H6)_n - H3'_{n-1}$

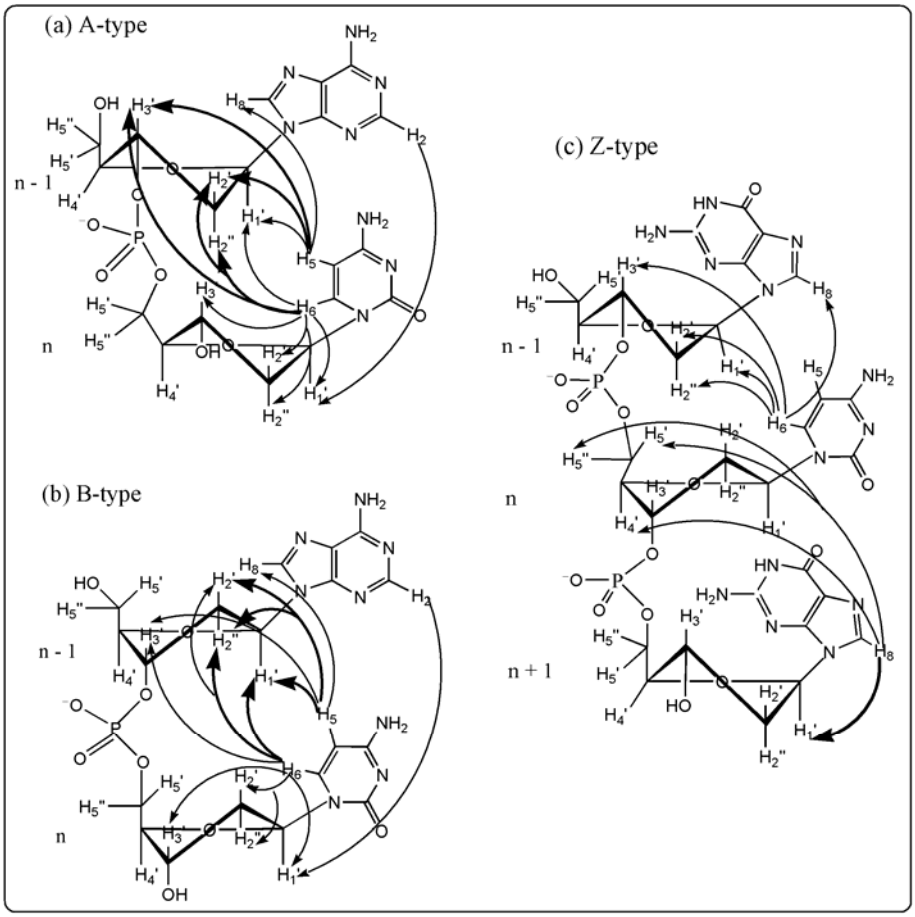


Figure 7. Interproton contacts less than approximately 5Å (marked by arrows) observed in nucleoside moieties of one of the strands of A-type (panel a), B-type (panel b) and Z- type (panel c) DNA/DNA duplex. Similar contacts are observed for RNA/RNA duplex. It should be noted that the distances involving exchangeable protons are omitted. The thick arrows in panel a, b and c are used to highlight the relatively shorter (within 3Å) interproton distances giving rise to stronger NOESY cross peaks.

/ $(H8/H6)_n - H3'_n$ cross-peak.

For both A- and B-type DNA/DNA or RNA/RNA duplex β^t and γ^+ is the preferred backbone conformation where the four bonds H4'-C4'-C5'-O5'-P are located in the same plane forming a W-shaped conformation. This is evident from the strong cross-peak intensity of P_n-$H4'_n$. Similarly if a detectable $^4J_{H2'P}$

coupling is observed in a DNA/DNA or RNA/RNA duplex it suggests the predominance of ε^{-} conformation in the duplex.

The NMR data (inter and intra proton distances and dihedral angles) can be used as constraints for determination of refined structures using molecular dynamics both at constant temperature (restrained molecular dynamics)[325-327] and by the temperature cycling protocol (simulated annealing)[326,327]. The protocols used for refinement of NMR structures are dependant on personal choices as well as trial and error. The final converged structures obtained are coherent with the experimental observation made by NMR and have a low RMSD (root-mean-square deviation) value between them. The structures obtained are not guaranteed to be at a global energy minima in the potential energy curve but are near to it.

6.3 Present work (Papers VII-IX)

Structural preorganization is very important for proper functioning of nucleic acids. The existing literature lacked in studies on structural aspects of single stranded nucleic acids which prompted us to investigate the helical pattern and stacking geometry of ssDNA as well as ssRNA. Extent of H bonding and stacking forces involved in duplex stabilization was also brought into light.

6.3.1 Observation of right handed helical pattern in ssDNA and ssRNA

The NOE footprint patterns in the NOESY spectra of hexameric ssDNA (**10d** Scheme 4) and ssRNA (**10c** Scheme 4) in aqueous solution have shown that they have right-handed helical structure. The presence of NOE connectivity [$(H6/H8)_n$ to $(H1', H2^{'/''}, H3')_{n-1}$] for both the ssRNA (**10c**) and ssDNA (**10d**) hexamers show that the single-stranded hexamers follow a right-handed helical conformation which is typical of the A or B-type double-stranded structures. The absence of $(H6/H8)_n$ to $(H5^{'/''})_{n-1}$ cross-peaks rules out the presence of Z-type conformation. In a comparison of NOESY footprints of hexameric ssDNA (**10d**) and ssRNA (**10c**) (see Table 4 and Figure 3 in **Paper VIII**) it has been found that the intra residue cross-peaks $(H6/H8)_n$–$(H3')_n$ and $(H5)_n$-$(H3')_n$, as well as the inter residue cross-peaks $(H6/H8)_n$–$(H3')_{(n-1)}$ and $(H5)_n$–$(H3')_{(n-1)}$, are significantly stronger in the ssRNA (**10c**) hexamer than in the ssDNA (**10d**) counterpart. In contrast, the interresidue $(H6/H8)_n$–$(H1')_{(n-1)}$

Table 3. Helical parameters and interproton distances that distinguish the A-type, B-type and Z-type DNA/RNA duplex and comparison with the ssDNA/ssRNA structures obtained in **Paper VIII**

Properties of DNA and RNA Helix		**A-type DNA/DNA or RNA/RNA**	**B-type DNA/DNA or RNA/RNA**	**Z-type DNA/DNA or RNA/RNA**	**ssDNA**	**ssRNA**
Helicity		Right handed	Right handed	Left handed	Right handed	Right handed
Base pair per repeat		1	1	2	1	1
Base pair per turn		11	10	12	-	-
Sugar conformation		North	South	South for C; North for G	South	North
Glycosidic bond		*anti*	*anti*	*anti* for C; *syn* for G	*anti*	*anti*
Minor groove	Width(Å)	11.0	5.7	4	-	-
	Depth(Å)	2.8	7.5	9	-	-
Major groove	Width(Å)	2.7	11.7	Convex	-	-
	Depth(Å)	13.5	8.5	Convex	-	-
Torsion angles	β	β^t	β^t	β^t	β^t	β^t
	γ	γ^+	γ^+	γ^+	γ^+	γ^+
	ε	ε^t	ε^t	ε^t	ε^-	ε^t
Interpro-ton distances	$H6/H8_n - H1'_{n-1}$	4.34	3.01	6.98^G $(4.8)^C$	S	S
	$H6/H8_n - H3'_{n-1}$	3.25	5.10	6.89^G $(3.55)^C$	M	S
	$H6/H8_n - H4'_{n-1}$	5.21	6.03	6.13^G $(4.76)^C$	M	M
	$H6/H8_n - H5'_{n-1}$	6.15/6.73	7.25/6.89	3.31^G $(6.08)^C$ / 4.66^G $(6.07)^C$	-/-	-/-
	$H5_n - H1'_{n-1}$	5.00	3.80	5.18^C	M	W
	$H5_n - H3'_{n-1}$	3.70	5.47	4.41^C	W	M
	$H2_n - H1'_{n-1}$	3.48	4.64	5.20	M	S
	$H5_n - H6/H8_{n-1}$	3.44	3.64	5.15^C	M	M

The helical parameters and cross-peak intensities categorized as strong (S, < 3Å) medium (M, 2.5-4 Å) or weak (4-5 Å) as obtained for hexameric ssDNA and ssRNA

are presented together for easy comparison. Superscript C indicates n is cytosine residue and superscript G indicates n is guanine residue. See reference 59-61 for other nucleobase containing duplexes that give Z-DNA

and $(H5)_n$ –$(H1')_{(n-1)}$ cross-peaks are stronger in ssDNA (**10d**) compared to ssRNA (**10c**). The distances involving the H3' proton is the main differences between the ribo and the deoxy hexamers (Figure 3E **Paper VIII**). The sugar proton $\Sigma^3 J_{H1'}$ couplings show that the sugar units of the hexameric ssDNA (**10d**) are predominantly in south conformation (C2'-*endo*), which is characteristic for B-type DNA duplex conformation, while the sugar units of the hexameric ssRNA (**10c**) are predominantly in north conformation (C3'-*endo*) (Table 1 **Paper VIII**), which is in turn associated with A-type conformation. Conformation for ssRNA (**10c**) was thus found to be predominantly right handed A-type helix and that of ssDNA (**10d**) was found to be right handed B-type helix.

6.3.2 NMR constraints and Molecular dynamics protocol for structural analysis of ssDNA and ssRNA

Starting ssDNA and ssRNA structures for molecular dynamics at constant temperature (298 K) were built from B-type antiparallel duplex for ssDNA hexamer and from A-type antiparallel duplex for ssRNA hexamer. The single stranded structures of hexameric DNA and RNA were solvated in water with Na^+ counterions used for neutralizing the phosphate charge. Molecular dynamics was run at 298 K (protocol described in **Paper VIII**) for 1.5 ns by introducing NMR constrains (Table 4) for both ssDNA and ssRNA. The simulation was stopped when RMSD and energy profile of the trajectory reached equilibrium at 298K.

Table 4. NMR constraints used in structure calculations

Type of NOE constrain	Number of constrains
Intranucleotide NOEs	95 for ssDNA; 70 for ssRNA
Internucleotide NOEs	46 for ssDNA; 41 for ssRNA
Dihedral restrains	16

Another set of simulations were performed with temperature cycling protocol (100 K to 450 K and back to 100K) with ssDNA and ssRNA (protocol in **Paper VIII**) to ensure that the starting structures can take a variety of conformations due to elevated temperatures. The starting ssDNA structure was the less preferred A-type and starting ssRNA structure was the less preferred B-type which were also solvated in a box filled with water molecules with added Na^+ ions. The NMR constrains used were same as in molecular dynam-

ics at constant temperature. The final set of structures obtained by simulated annealing was the same as obtained by constrained molecular dynamics and the RMSDs were also comparable (<1.8Å for ssDNA and <1.7Å for ssRNA)

6.3.3 Difference in stacking geometry of ssDNA vs. ssRNA as observed from NMR constrained MD simulation

pH titration of ssRNA (**10c**) showed that the δH8A is deshielded with pH while δH2A remains non-responsive. In ssDNA (**10d**) on the other hand δH2A is deshielded with pH and δH8A remains non-responsive. This hints at different orientation of the base-base stacking in the ground state in ssRNA (**10c**) compared to ssDNA (**10d**). The NMR- constrained molecular dynamics showed that overlap of neighboring nucleobase of ssRNA as well as ssDNA is guided by attractive electrostatics between the relatively electron-deficient pyrimidine and the relatively electron-rich imidazole moieties at each dinucleotide step. The difference between ssRNA and ssDNA lies in the relative directionality of this interaction. The MD snapshots zooms highlighting the stacking geometries of each dinucleotide step of the hexameric ssDNA (**10d**) and ssRNA (**10c**) (Figure 6, panels C and D **Paper VIII**) show that the imidazole stacks above the pyrimidine in the 5' to 3' direction in ssDNA (**10d**) while, in contradistinction, the pyrimidine stacks above the imidazole in the 5' to 3' direction in ssRNA (**10c**). In ssDNA, the H2 protons are completely overlapped by the shielding cone of the base from the 5'- side (*n* - 1), while the H2 protons of ssRNA (**10c**) are mainly extruded and only partly shielded by the base from the 3'- side (*n* + 1). The situation is reversed for the H8 protons. In ssDNA, (**10d**) the H8 protons are extruded, while they receive full shielding by the base to the 5'-side (*n* - 1) in ssRNA (**10c**). This means that the negative π-surface of one pyrimidine ring is near to the relatively positive imino and amino hydrogens normally involved in base pairing of the *n* + 1 and *n* - 1 neighboring base for DNA and RNA, respectively.

The chemical shifts of the ssDNA (**10d**) and ssRNA (**10c**) base protons have been used to confirm the stacking pattern found in the MD simulations (Figure 6 **Paper VIII**). The displacement of the H6, H8, and H2 protons in the *X*-*Y* plane, relative to the closest ring atom of the neighboring bases (*n* + 1 and *n* - 1) have been measured and compared with the chemical shift shifts (Δδ) obtained both from oligomerization (oligomerization shift = $\Delta\delta_{monomer-oligo}$) and destacking (destacking shift = $\Delta\delta_{363K\ -\ 298K}$) processes. Plots of oligomerization shift and destacking shift versus the displacement in the *X*-*Y* plane (Figure 7 **Paper VIII**) for each proton from the nearest edge of the neighboring nucleobases measured in both 3' and 5' directions showed high correlation of 0.83 and 0.82 respectively. It is noteworthy that the correlation is completely lost when the coordinates of ssDNA (**10d**) and ssRNA (**10c**) are

reversed. This shows the difference in stacking orientations in ssDNA (**10d**) compared to ssRNA (**10c**).

6.3.4 Extent of stacking versus base pairing forces involved in relative stability of RNA-RNA compared to DNA-DNA duplex, estimated from the pK_a calculation of the model mononucleotides

pK_a values of nucleobases in nucleoside 3'-ethylphosphates (**3a** – **3e** for ribo and **3f** – **3j** for deoxy, Scheme 1) as well as nucleoside 3',5'-bis-ethylphosphates (**5a** – **5e** for ribo and **5f** – **5j** for deoxy, Scheme 1) in both 2'-deoxy and ribo series obtained from pH-dependant ^{1}H NMR was used to study relative stabilities of RNA-RNA and DNA-DNA duplex. The nucleoside 3',5'-bis-ethylphosphates (**5a** – **5e** for ribo and **5f** – **5j** for deoxy, Scheme 1) were used to mimic the internucleotidic monomeric unit where the stacking contribution normally found in DNA and RNA duplexes is completely absent. All monomeric DNA nucleobases are found to be more basic than the corresponding RNA nucleobases by ~0.2 pK_a unit due to the presence of electron withdrawing effect of 2'-OH group in RNA. It was found that from the pK_a of the monomeric nucleotide blocks, ΔpK_a values can be used to understand the relative contribution of base-pairing vis-à-vis stacking in oligomeric duplexes. Thus the ΔpK_a values for the G-C base-pairing with model ribo pair (**5a**/**5c**) and deoxy pair (**5f**/**5h**) are 5.04 and 5.24 respectively (Table 2A in **Paper IX**). Similarly, the ΔpK_a for the A-U/T base-pairing with model ribo pair (**5b**/**5d**) and deoxy pair (**5g**/**5i**) are 5.53 and 6.29 respectively. Comparison of these ΔpK_a values of G-C and A-T/U base-pairing shows that the ribo base-pairing is stronger (*i.e.* ΔpK_a is less) than the corresponding deoxy base pairing. Also the ΔpK_a is considerably less for the G-C base-pairing in both ribo and deoxy series than those for the A-U/T base-pairing (Table 2A in **Paper X**) which is consistent with the fact that the former is stronger than the latter.[2,3]

The $\Delta G°_{pKa}$ of the donor and acceptor of the base pair in duplexes enables a qualitative dissection of the relative strength of the base-pairing ($[\Delta G^o_{bp}]_{RR–DD}$) and stacking ($[\Delta G^o_{stacking}]_{RR–DD}$) in the RNA-RNA over the DNA-DNA duplexes. Linear plots of $[\Delta G^o_{bp}]_{RR–DD}$ and $[\Delta G^o_{stacking}]_{RR–DD}$ as a function of % A-T/U bp content (Figure 3 in **Paper X**) show opposite slopes. This shows that with increasing content of A-T/U base pairs the base-pairing stability of DNA-DNA duplex weakens over the corresponding RNA-RNA duplexes ($[\Delta G^o_{bp}]_{RR–DD}$), while the strength of stacking ($[\Delta G^o_{stacking}]_{RR–DD}$) 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 interaction[237] between the 5-methyl group of T with the nearest-neighbor A in the AT rich sequence.

6.4 Implications

The above study includes elucidation of structural characteristic of single stranded nucleic acids, the directionalities involved in base-base stacking and factors governing stability of duplexes. This can help us in understanding the intermolecular forces involved in homo or hetero duplex formation, triplex formation or interactions with specific ligands. The above knowledge can be used in understanding folding through secondary structure formation in single stranded nucleic acids. This knowledge can also be used in designing of various therapeutic agents.

Acknowledgements

This work has been done in the Department of Bioorganic Chemistry, Uppsala University (from December 1999 to April 2006). In the words scribbled below, I take the opportunity to thank everybody who helped and supported me in this journey.

To begin with, I express my sincere gratitude to my supervisor Professor Jyoti Chattopadhyaya for accepting and supporting me as a PhD student in his well-equipped laboratory, introducing me to the field of nucleic acid chemistry and providing me guidance throughout these years.

I sincerely thank Dimitri Ossipov, Melker Holmer, Catherine Petit, Edouard Zamaratski, Mrinal Kanti Kundu, Gopal Dutta, S. Karthick Babu, Dimitri Honcharenko, Dimitri Kulik, Xiaoguang Luo, Pei Wen, Brian Lawrance, Puneet Sreevastava, Malgorzata Bogucka, Petr Naus, Thirumari Mariappan, Mohitosh Maiti, Ayan Samanta, Hu Wenbin, Dimitri Ovcharenko, for maintaining a friendly collaborative atmosphere in this department over the years.

My special thanks are due to:

Dr. Tatiana Maltseva for providing me initial hands-on training with high field NMR spectrometers.

Dr. Andras Földesi for offering me sincere help whenever I approached him with a problem. Thak you for thoroughly checking my thesis.

Dr. Oleksandr Plashkevych for promt help regarding computer related problems in the laboratory and for his comments on my thesis.

Dr. Steve Graham for his critical comments on my thesis.

Dr. Nariman Amirkhanov, Anna Trifonova, Pradeep Cheruku and Zhou Chuanzheng for taking the pains of synthesizing and purifying a number of compounds used in our study.

Jharna Barman for synthesizing the single stranded DNA and RNA oligomers as well as helping me generously in and outside the lab.

Omen P. Varghese for his good friendship.

Dr. Johan Issakson for being a valuable co-worker, Vimal Pathmasiri for his always ready to help attitudes and Subhrangshu Chtterjee for NMR experiments with decomposable RNAs and related works.

Irina Velikyan for being nice and friendly colleague when I first started working in this laboratory.

Jaana Öst and Rabab Elkarib for their kind assistance and competent secretarial work. I also wish to thank all staffs from BMC-biblioteket and from BMC administration for their kind help.

Thanks to my teachers in school and university for helping me in all my endauvers. I would specially like to thank Professor R. R. Rao (Late) and Professor Aloka Rao for their encouragement to persue research in chemitry.
I would like to thank all my friends and the growing Indian community in Uppsala which never make me feel so far away from home.
Special thanks to my friends Dr. Bharat Bhikaji; Dr. Pradeep Kumar; Dr. Rituparna Mukherjee; Drs. Biplab and Suparna Sanyal; Somali and Dr. Sabyasachi Sanyal; Drs. Kaushik and Pritha Mahato for being kind and affectionate friends.
Thanks a due to Dr. Santanu Dasgupta and Dr. Lena Åslund for being affectionate and specially Lena for helping me out with the Swedish summary of this thesis in the last moment.
At the end I express my heart felt thanks to my dear family. First of all I thank my parents and my brother for creating the best environment in a family and sharing all my joys and sorrows in life. I specially thank my mother for comming all the way to Sweden to take care of my little son (Protit) as I have been preparing my thesis. I would also like to thank my father-in-law and mother-in-law for being so supportive and concerned about me and helping me out with Protit. I have no words to appreciate the love and support I got from my husband both as a colleauge and partner in life. I highly appreaciate his cool nature and patient attitudes. Last but not the least I thank my son Protit for always cheering me up in home with his cute smiles.

Summary in Swedish

Denna avhandling är baserad på nio vetenskapliga publikationer (**I-IX**) som sammanfattar mina studier på nukleinsyror (DNA/RNA), de molekyler som utgör vår arvsmassa och som är uppbyggda av flexibla sockerenheter, en ryggrad av polyelektrolytisk fosfodiester och nukleobaser. De är, med hjälp av andra faktorer, involverade i fundamentala biologiska funktioner såsom bevarandet av den genetiska koden på DNA-nivå och i katalytiska reaktioner genom RNA. Olika egenskaper hos nukleinsyror för bindning av ligander och igenkänning medieras genom inter- och intramolekylära vätebindningar och stackningsinteraktioner. I detta arbete användes det pH-beroende kemiska skiftet, pK_a värden av 2'-OH gruppen eller nukleobaserna, alkalisk hydrolys av internukleotid fosfodiesterbindningar och analys av 'NOESY footprint' studier tillsammans med NMR-begränsad molekylärdynamisk simulering, som verktyg för att undersöka det fysikalkemiska beteendet hos olika nukleinsyresekvenser med syfte att förstå reaktiviteten och krafterna involverade i deras själv-sammansättande. **Arbete I-II** visade att joniseringen av 2'-OH gruppen är nukleobas-beroende. **Arbete III** visade att de kemiska egenskaperna hos internukleotidfosfat är icke-identiska i RNA jämfört med dessa i DNA. **Arbete IV-VI** visade att variabla intramolekylära elektrostatiska interaktioner mellan elektroniskt kopplade närmaste grann-nukleobaser i ett enkelsträngat RNA kan modulera deras respektive pseudoaromatiska egenskaper. Detta kan resultera i bildandet av en unik uppsättning av aglykoner med unika egenskaper beroende på tendensen och geometrin av närmaste grannbas interaktionen. **Arbete VII** visade att kors-moduleringen av den pseudoaromatiska karaktären hos nukleobasen genom närmaste grannbas-interaktionen till naturen är sekvensberoende och denna effekt kan överföras längs med oligonukleotidkedjan. **Arbete VIII** visade att det purinrika hexameriska enkelsträngade DNAt och RNAt behåller den högervridna helixstrukturen (B-typ i enkelsträngat DNA och A-typ i enkelsträngat RNA) i enkelsträngad form, även i frånvaro av intermolekylär vätebindning. Riktningen av stackningsgeometrin skiljer sig emellertid i enkelsträngat DNA jämfört med enkelsträngat RNA. I enkelsträngat DNA stackas det relativt elektronrika imidazolet ovanför den elektronfattiga pyrimidinet i 5' till 3' riktning, i klar motsats till enkelsträngat RNA där pyrimidinet stackas ovanför imidazolet i 5' till 3' riktning. **Arbete IX** visade att pK_a värden hos nukleobaserna i monomeriska nukleotider kan användas för att visa att en RNA-RNA duplex är mer stabil än en DNA-DNA duplex. Dissektion av den relativa styrkan i basparning och

stackning visade att bidraget från basparning i relation till stackning i en RNA-RNA duplex jämfört med motsvarande DNA-DNA duplex, minskar med ökande innehåll av A-T/U baspar i en sekvens.

References

1. Limbach, P. A.; Crain, P. F.; Mc closky, J. A.; *Nucleic Acids Res.* **1994**, *22*, 2183.
2. Saenger, W. *Principles of Nucleic Acid Chemistry*; Springer-Verlag, New York, 1984
3. Bloomfield, V. A.; Crothers, D. M.; Tinoco, I. *Nucleic Acids: Structures, Properties and Functions*; University Science Books, Sausalito, California., 1999
4. Blackburn, G. M.; Gait, M. J. *Nucleic Acids in Chemistry and Biology*; Oxford University Press, NewYork, 1996.
5. Thibaudeau, C.; Acharya, P.; Chattopadhyaya, J. *Stereoelectronic Effects in Nucleosides and Nucleotides and their Structural Implications*, (ISBN 91-506-1351-0), Department of Bioorganic Chemistry, Uppsala University Press (jyoti@boc.uu.se), Sweden, 2005.
6. Carreira, L. A.; Jiang, G. J.; Person, W. B.; Willis, J. N. J. *J. Chem. Phys.* **1972**, *56*, 1440.
7. Kilpatrick, J. E.; Pitzer, K. S.; Spitzer, R. *J. Am. Chem. Soc.* **1947**, 69, 2483
8. Strauss, H. L. *Ann. Rev. Phys. Chem.* **1983**, *34*, 301.
9. Geise, H. J.; Altona, C.; Romers, C. *Tetrahedron Lett.* **1967**, *15*, 1383.
10. Altona, C.; Geise, H. J.; Romers, C. *Tetrahedron* **1968**, *24*, 13.
11. Altona, C.; Sundaralingam, M. *J. Am. Chem. Soc.* **1972**, *94*, 8205.
12. Altona, C.; Sundaralingam, M. *J. Am. Chem. Soc.* **1973**, *95*, 2333.
13. de Leeuw, H. P. M.; Haasnoot, C. A. G.; Altona, C. *Isr. J. Chem.* **1980**, *20*, 108.
14. Röder, O.; Ludemann, H. D.; vonGoldammer, E. Eur. J. Biochem. **1975**, *53*, 517.
15. Houseknecht, J. B.; Lowary, T. L. *J. Org. Chem.* **2002,** *67*, 4150.
16. Feigon, J.; Wang, A. H. J.; van der Marel, G. A.; van Boom, J. H.; Rich, A. *Nucleic Acids Res.* **1984**, *12*, 1243.
17. Tran-Dihn, S.; Taboury, J.; Neumann, J. M.; Huynh-Dinh, T.; Genissel, B.; Langlois d'Estaintot, B.; Igolen, J. *Biochemistry*, **1984**, *23*, 1362.
18. Davis, P. W.; Hall, K.; Cruz, P.; Tinoco, Jr.; Neilson, T. *Nucleic Acids Res*. **1986**, *14*, 1279.
19. Davis, P. W.; Adamiak, R. W.; Tinoco, I. J. *Biopolymers* **1990**, *29*, 109.
20. Agback, P.; Sandström, A.; Yamakage, S. I.; Sund, C.; Glemarec, C.; Chattopadhyaya, J. *J. Biochem. Biophys. Meth.* **1993**, *27*, 229.
21. Agback, P.; Glemarec, C.; Yin, L.; Sandström, A.; Plavec, J.; Sund, C.; Yamakage, S. I.; Viswanadhm, G.: Rousse, B.; Puri, N.; Chattopadhyaya, J. *Tetraherdon Lett.* **1993**, *34*, 3929.
22. Miha, P.; Cmugelj, M.; Stimac, A.; Kobe, J.; Plavec, J. *J. Chem. Soc. Perkin Trans*. **2001**, *28*, 1433.
23. Thibaudeau, C.;Plavec, J.; Chattopadhyaya, J. *J. Org. Chem.* **1996**, *61*, 266.
24. Herdewijn, P. *Biochem. Biophys. Acta.* **1999**, *489*, 167.
25. P. I. Pradeepkumar *PhD. Thesis, (ISBN 91-554-5957-9) Department of Bioorganic Chemistry, Uppsala University, Sweden* 2004.

26. Plavec, J.; Thibaudeau, C.; Chattopadhyaya, J. *Pure Appl. Chem.* **1996**, 68, 2137.
27. Guschlbauer, W.;Jankowski, K. *Nucleic Acids Res.* **1980**, 8, 1421.
28. Watson, J. D.; Crick, F. H. C. *Nature* **1953**, *171*, 737.
29. Kennard, O.; Hunter, W. N. *Q. Rev. Biophys.* **1989**, *22*, 327.
30. Pearlman, D. A.; Kim, S. G. *J. Mol. Biol.* **1990**, *211*, 171.
31. Sponer, J.; Leszczynski, J.; Hobza, P. *Biopolymers*, **2001**, *61*, 3
32. Sponer, J.; Leszczynski, J.; Hobza, P. *J. Phys. Chem*, **2001**, *100*, 1965.
33. Marion, M. N.; Xiangming, K.; LiWang, A. J. Am. Chem. Soc. 2005, 127, 17974.
34. Ts'o, P. O. P. Bases, Nucleosides and Nucleotides, in *Basic Principles in Nucleic Acid Chemistry*, Ts'o, P. O. P., Ed., Vol. I, Academic, New York. 1974.
35. Burkard, M. E.; Kierzek, R.; Turner, D. H. *J. Mol. Biol.* **1999**, *290*, 967.
36. Kim, J.; Walter, A. E.; Turner, D. H. *Biochemistry* **1996**, *35*, 13753.
37. Bommarito, S.; Peyret, N.; SantaLucia, J., Jr. *Nucleic Acids Res.* **2000**, *28*, 1929.
38. Rosemeyer, H.; Seela, F. *J. Chem Soc., Perkin Trans. 2*, **2002**, 746.
39. Ohmichi, T.; Nakano, S-i.; Miyoshi, D.; Sugimoto, N. *J. Am. Chem. Soc*. **2002**, *124*, 10367.
40. Falk, M.; Hartman, K. A. Jr.; Lord, R. C. *J. Am. Chem. Soc.* **1963**, *85*, 387.
41. Wiberly, B.; Varani, G.; Tinoco, I. J. *Biochemistry* **1993**, *32*, 1078.
42. Hoogsteen, K. . *Acta Cryst.*. **1963**, *16*, 907.
43. Kim, S. G. *Prog. Nucl. Acid Res. Mol. Biol.* **1976**, *17*, 181.
44. Sussman, J. L.; Holbrook, S. R.; Wade Warrant, R.; Church, G. M.; Kim, S. G. *J. Mol. Biol.* **1978**, *123*, 607.
45. Holbrook, S. R.; Cheong, C.; Tinico, I. Jr.; Kim, S. H. *Nature*, **1991**, *353*, 579.
46. Brodsky, A. S.; Erlache, H. A.; Williamson, J. R. *Nucleic Acids Res.* **1998**, *26*, 1991
47. Dulbecco,R.; Vogt, M. *Proc. Natl. Acad. Sci. U.S.A.***1963**, *50*, 236.
48. Weil,R.; Vinograd, J. *Proc. Natl. Acad. Sci. U.S.A.***1963**, *50*, 730.
49. Murchie, A. I. H.; Lilley, D. M. J. *Methods Enzymol.* **1992**, *211*, 158.
50. Kallenbach, N.; Zhong, M. *Curr. Opin. Struct. Biol.* **1994**, *4*, 365.
51. Schultze, P.; Macaya, R. F.; Feigon, J. *J. Mol. Biol*. **1994**, *235*, 1532.
52. Nagai, K.; Mattaj, I. W. *RNA-Protein Interactions*, IRL Press at Oxford University Press, 1994.
53. Leslie, A. G. W.; Arnott, S.; Chandrasekaran, R.; Ratliff, R. L. *J. Mol. Biol.* **1980** 143, 49.
54. Wang, A. H. J.; Quigley, G. J.; Kolpak, F. J.; Crawford, J. L.; van Boom, J. H.; van der Marel, G.; Rich, A. *Nature* **1979**, *282*, 680.
55. Dickerson, R. E. *Methods Enzymol.* **1992**, *211*, 67.
56. Trulson, M. O.; Cruz, P.; Puglisi, J. D.; Tinoco, I.; Mathies, R. A. *Biochemistry* **1987**, *26*, 8624.
57. Pohl, F. M.; Jovin, T. M. *J. Mol. Biol.* **1972**, *67*, 375.
58. Behe, M.; Felsenfeld, G. *Proc. Natl. Acad. Sci. USA* **1981**, *78*, 1619.
59. Aiken, J. M.; Miller, F. D.; Hagen, F.; McKenzi, D. I.; Krawetz, S. A.; van de Sande, J. H.; Rattner, J. B.; Dixon, G. H. *Biochemistry*, **1985**, *24*, 6268.
60. Haniford, D. B.; Pulleyblank, D. E. Nature **1983**, *302*, 632.
61. Nordheim, A.; Rich, A. *Proc. Natl. Acad. Sci. U.S.A* **1983**, *80*, 1821.
62. Dickerson, R. E. *J. Mol. Biol.* **1998**, *205*, 787.
63. Patel, D. J.; Suri, A. K. *Rev. Mol. Biotechnol.* **2000**, 74, 39.
64. Nimjee, S. M.; Rusconi,, C.; Sullenger, B. A. *Annu. Rev. Med.* **2005**, *56*, 555,
65. Gilbert, D. E.; Feigon, J. *Curr. Opin. Struct. Biol.* **1999**, *9*, 305.

66. Gesteland, R. F.; Cech, T. R.; Atkins, J. F. *The RNA World*, Cold Spring Harbour Laboratory Press, Cold Spring Harbour, New York.1998.
67. Fersht, A. R. *Structure and Mechanism in Protein Sciences*, W. H. Freeman and Co, San Francisco, 1999.
68. Fasman, G. D. Handbook of Biochemistry and Molecular Biology, 3rd ed., Nucleic Acids, Vol. 1, CRC press, Cleveland, OH. 1975.
69. Chamberlin, S.; Merino, E. J.; Weeks, K. M. *Proc. Natl. Acad. Sci.* **2002**, *99*, 14688.
70. Kumler, W. D.; Eiler, J. J. *J. Am. Chem. Soc.* **1943**, *65*, 2355.
71. Olsthoorn, S. M.; Bostelaar, L. J.; DeRooij, J. F. M.; VanBoom, J. H.; Altona, C. *Eur. J. Biochem.* **1981**, *115*, 309.
72. Freier, S. M.; Hill, K. O.; Dewey, T. G.; Marky, L. A.; Breslauer, K. J.; Turner, D. H. *Biochemistry* **1981**, *20*, 1419.
73. Xia, T.; SantaLucia, J. Jr.; Burkard, M. E.; Kierzek, R.; Schroeder, S. J.; Jiao, X.; Cox, C.; Turner, D. H. *Biochemistry* **1998**, *37*, 14719.
74. Williams, A. P.; Longfellow, C. E.; Freier, S. M.; Kierzek, R.; Turner, D. H. Biochemistry, **1989**, *28*, 4283.
75. Joule, J. A.; Mills, K. *Heterocyclic Chemistry*. Fourth Edition. Blackwell Science, 2000.
76. Hunter, C. A.; Sanders, J. K.M. *J. Am. Chem.Soc.* **1990**, *112*, 5525.
77. Fohrer, J.; Henning, M.; Carlomagno, T. *J. Mol. Biol.* **2006**, *356*, 280.
78. Popenda, M.; Milecki, J.; Adamiak, R. W. *Nucleic. Acids. Res.* **2004**, *32*, 4044
79. DeRose, V. J. *Curr. Opin. Struc. Biol.* **2003**, *13*, 317.
80. Fierro-Monti, I.; Mathews, M. B. *Trends. Biomol. Sc.* **2000**, 241.
81. Yoshida, A.; Shan, S.; Herschlag, D.; Piccirilli, J. A. *Chem. Biol.* **2000**, *7*, 85
82. Shan, S.; Yoshida, A.; Sengen, S.; Piccirilli, J. A.; Herschlag, D. *P.N.A.S* **1999**, *96*, 12299
83. Williamson, J. R.; Fedor, M. J. *Nature Reviews Mol. Cell. Boil.* **2005**, *6*, 399
84. Takagi, Y.; Ikeda, Y.; Taira, K. *Top. Curr. Chem.* **2004**, *232*, 213.
85. Bevilacqua, P. C.; Brown, T. S.; Nakano, S.; Yajima, R. *Biopolymers* **2004**, *73*, 90.
86. Lilley, D. *Trends. Biochem. Sci.* **2003**, *9*, 495.
87. DeRose, V. J. *Chem. Biol.* **2002**, *9*, 961.
88. Schroeder, R.; Grossberger, R.; Pichler, A.; Waldsich, C. *Curr. Opin. Struct. Biol.* **2002**, *12*, 296.
89. Fedor, M. J. *Curr. Opin. Struct. Biol.* **2002**, *12*, 289.
90. Doudna, J. A. and Cech, T. R. *Nature* **2002**, *418*, 222.
91. Hermann, D. *Bioorganic Chemistry-A Chemical approach to Enzyme Action.* 3rd Edn. Springer-Verlag, 1996
92. Walsh, C. *Enzymatic Reaction Mechamism*; New York, W. H. Freeman, 1979.
93. Gerlt, J. A. *The Nucleobases*; Cold Spring Harbour Laboratory Press, 1993.
94. Jack, A.; Ladner, J. E.; Rhodes, D.; Brown, R. S.; Klug, A. *J. Mol. Biol.* **1977**, *111*, 315.
95. Hermann, T.; Patel, D. J. *J. Mol. Biol.* **1999**, *294*, 829.
96. Tamura, M.; Holbrook, S. R. *J. Mol. Biol.* **2002**, *320*, 455.
97. Quigley, G. J.; Seeman, N. C.; Wang, A. H.; Suddarth, F. L.; Rich, A. *Nucleic Acids Res.* **1975**, *2*, 2329 .
98. Quigley, G. J.; Wang, A. H.; Seeman, N. C.; Suddarth, F. L.; Rich, A.; Sussman, J. L.; Kim, S. H. *Proc. Natl. Acad. Sci. U.S.A* **1975**, *72*, 4866 .
99. Quigley, G. J.; Rich, A. *Science* **1976**, *194*, 796
100. Moore, P.B. *Annu. Rev. Biochem.* **1999**, *68*, 287

101. Nissen, P.; Ippolito, J.A.; Ban, N.; Moore, P.B.; Steitz, T. A. *Proc. Natl. Acad. Sci. U.S.A* **2001**, *98*, 4899
102. Cate, J. H.; Gooding, A.R.; Podell, E.; Zhou, K.; Golden, B. L.; Kundrot, C. E.; Cech, T. R.; Doudna, J. A. *Science* **1996**, *273*, 1678
103. Ferre-d Amare, A. R.; Zhou, K.; Doudna, J. A. *Nature* **1998**, *395*, 567
104. Rupert, P. B.; Ferre-d Amare, A. R. *Nature* **2001**, *410*, 780
105. Auffinger, P.; Westhof, E. *RNA* **2001**, *7*, 334
106. Williams, J. D.; Boots, J. L.; Hall, K. B. *RNA* **2001**, *7*, 44
107. SantaLucia, J. Jr.; Kierzek, R.; Turner, D. H. *Science* **1992**, *256*, 217
108. Williams, J. D.; Hall, R. *J. Mol. Biol* **2000**, *297*, 1045
109. Frederic, H. T. A.; Varani, G. *J. Mol. Biol* **1995**, *250*, 333
110. Varani, G. *Annu. Rev. Biophys. Biomol. Struct.* **1995**, *24*, 379
111. Millar, J. L.; Kollman, P. A. *Annu. Rev. Biophys. Biomol. Struct.* **1995**, *24*, 379
112. Miller, J. L.; Kollman, P. A. *J. Mol. Biol* **1997**, *270*, 436
113. Pleiss, J. A.; Ullenbeck, O. C. *J. Mol. Biol* **2001**, *308*, 895
114. Hou, Y. M.; Zhang, X.; Holland, J. A.; Davis, D.R. *Nucleic Acids Res.* **2001**, *29*, 976
115. Jiang, F.; Kumar, R. A.; Jones, R. A.; Patel, D. J. *Nature* **1996**, *382*, 183
116. Kisreboom, L. A.; Svard, S. G. *Nucleic. Acids. Res.* **1992**, *20*, 425
117. Frank, D. N.; Pace, N. R. *Annu. Rev. Biochem.* **1998**, *67*, 153
118. Bevilacqua, P. C.; Turner, D. H. *Biochemistry.* **1991**, *30*, 10632
119. Pyle, A. M.; Cech, T. R. *Annu. Rev. Biochem.* **1991**, *67*, 153
120. Strobel, S. A.; Cech, T. R. *Biochemistry.* **1993**, *32*, 13593
121. Herschlag, D.; Eckstein, F.; Cech, T. R. *Biochemistry.* **1993**, *32*, 8299
122. Narlikar, G. J.; Herschlag, D. *Biochemistry.* **1998**, *37*, 9902
123. Muller, U. F.; Bartel, D. P. *Chemistry and Biol.* **2003**, *10*, 799
124. St. Johnston, F.; Brown, N. H.; Gall, J. G.; Jantsch, M. *P.N.A.S* **1992**, *89*, 10979
125. Bevilacqua, P. C.; Cech, T. R. *Biochemistry.* **1996**, *35*, 9983
126. Ravindranathan, S.; Butcher, S. E.; Feigon, J. *Biochemistry* **2000**, *39*, 16026.
127. Cai, Z.; Tinoco, I. Jr. *Biochemistry* **1996**, *35*, 6026.
128. Nakano, S.; Chadelavada, D. M.; Bevilacqua, P. C. *Science*, **2000**, *287*, 1493.
129. Lupták, A.; Ferré-D'Amaré, A. R.; Zhou, K.; Zilm, K. W.; Doudna, J. A. *J. Am. Chem. Soc*. **2001**, *123*, 8447.
130. Izatt, R. M.; Hansen, L. D.; Rytting, J. H.; Christensen, J. J. *J. Am. Chem Soc.* **1965**, *87*, 2760
131. Levene, P. A.; Bass, L. W.; Simms, H. S. *J. Biol. Chem.***1926***, 70,* 229
132. Usher, D. A.; Richardson, D. I. Jr.; Oakenfull, D. G. *J. Am. Chem Soc.* **1970**, *92*, 4699
133. Lyne, P. D.; Karplus, M. *J. Am. Chem.Soc.***2000***, 122,* 166
134. Izatt, R. M.; Rytting, J. H.; Hansen, L. D.; Christensen, J. J. *J. Am. Chem Soc.* **1966**, *88*, 2641
135. Christensen, J. J.; Rytting, J. H.; Izatt, R. M. *J. Phy. Chem.* **1967**, *71*, 2700.
136. Kolke, T.; Inoue, Y. *Chem. Lett.* **1972**, 569.
137. Bock, R. M. *Methods. Enzymol.* **1967**, *XIIA*, 218.
138. Liu, X.; Reese, C. B. Tetrahedron Lett. **1995**, *36*, 3413.
139. Olvanen, M.; Ora, M.; Almer, H.; Stromberg; Lönnberg, H. *J. Org. Chem* **1995**, *60*, 5620.
140. Weinstein, L. B.; Earnshaw, D. J.; Cosstick, R.; Cech, T. R. *J. Am. Chem Soc.* **1996**, *118*, 10341
141. Järvinen, P.; Oivanen, A.; Lönnberg, H. *J. Org. Chem* **1991**, *56*, 5396
142. Li, Y.; Breaker, R. R. *J. Am. Chem.Soc.***1999***, 121,* 5364

143. Albert, A.; Serjeant, E. P. *The determination of Ionization constants* London, 1971
144. Perrin, D. D. Aust. J. Chem. **1964**, 17, 484.
145. The pK_a of the 2'-OH group in nucleos(t)ides reported in Paper I of this thesis and the pK_a of the internucleotidic 2'-OH in dinucloetide monophosphates reported in Paper II of this thesis has been recalculated because of an error in defining the pH values. Since we have used titration with a known strength of H_2SO_4 for determining the pH from 13.0 to 14.0 in case of nucloes(t)ides and pH 11.1 to 14 of dinucleotide monophosphates, without any readout from the electrode, it is not needed to reduce 0.4 pH units from the pD values to obtain the corresponding pH values. The Tables I and II in this thesis contains the corrected pK_a values.
146. Lönnberg, T.; Lönnberg, H. *Curr. Opin. Chem. Biol.* **2005**, *9*, 665.
147. Perreault, D.M.; Anslyn, E.V. *Angew. Chem. Int. Ed. Engl.* **1997**, *36*, 432.
148. Oivanen, M.; Kuusela, S.; Lönnberg, H. *Chem. Rev.* **1998**, *98*, 961.
149. Soukup, G.A.; Breaker, R.R. *RNA* **1999**, *5*, 1308.
150. Lane, B.G.; Butler, G.C. *Biochim. Biophys. Acta.* **1959**, *33*, 281
151. Kierzek, R. *Nucleic Acid Res.* **1992**, *20*, 5079.
152. Kierzek, R. *Methods in Enzymology*, **2001**, *541*, 657.
153. Zagorowska, I.; Mikkola, S.; Lonnberg, H. *Helvetica Chimica Acta*, **1999**, *82*, 2105.
154. Williams, K. P, Ciafré; Tocchini-valentine, G. P, *EMBO J,* **1995**, *14*, 4551-4557.
155. Bibillo, A., Figlerowicz, M., Kierzek, R. *Nucleic Acid Res.* **1999**, *27*, 3931.
156. Brown D.M.; Magrath D. I., Neilson A. H. *Nature* **1956,** *177*, 1124
157. Oivanen M., Schnell R., Pfleiderer W., Lönnberg H. *J. Org. Chem.* **1991**, *56*, 3623
158. Jarvinen P., Oivanen, M., Lonnberg, H. *J. Org. Chem.* **1991**, *56*, 5396.
159. Kaukinen, U.; Lyytikäinen, S.; Mikkola, S.; Lönnberg, H. *Nucleic Acid Res.* **2002**, *30*, 468.
160. Kaukinen, U.; Venäläinen, T.; Lönnberg, H.; Peräkylä, M. *Org. Biomol. Chem.* **2003**, *1*, 2439.
161. Kaukinen, U.; Lönnberg, H.; Peräkylä, M. *Org. Biomol. Chem.* **2004**, *2*, 66.
162. Kawamura, K. *Bull. Chem. Soc. Jpn.* **2003**, *76*, 153.
163. Bibillo, A.; Figlrowicz, M.; Ziomek, K.; Kierzek, R. Nucl. Nucl.**2000**, *19*, 977.
164. Hertel, K. J.; Peracchi, A.; Uhlenbeck, O. C.; Herschlag, D. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 8497.
165. Narlikar, G. J.; Herschlag, D. *Biochemistry* **1998**, *37*, 9902.
166. Zhang, L.; Doudna, J. A. *Science* **2002**, *295*, 2084.
167. Ferre-D Amare, A. R.; Zhou, K.; Doudna, J. A. *Nature* **1998**, *395*, 567.
168. Ke, A.; Zhou, K.; Ding, F.; Cate, J. H. D.; Doudna, J. A. *Nature* **2004**, *429*, 201.
169. Rupert, P. B.; Ferre-D Amare, A. R. *Nature* **2001**, *410*, 780.
170. Rupert, P. B.; Massey A. P.; Sigurdsson S. T.; Ferre-D Amare, A. R. *Science* **2002**, *298*, 1421.
171. Niittymäki, T.; Lönnberg, H. *Org. Biomol. Chem.* **2006**, *4*, 15.
172. Lilley, D. M. J. *RNA* **2004**, *10*, 151.
173. Rould, M.A.; Perona, J.J.; Soll, D.; Steitz, T.A. *Science* **1989**, *246*, 1135
174. Antson, A.A. *Curr. Opin. Struct. Biol.* **2000**, *10*, 87.
175. Handa, N.; Nureki, O.; Kuriomoto, K.; Kim, I.; Sakamoto, H.; Shimura, Y.; Muto, Y.; Yokoyama, S. *Nature* **1999**, *398*, 579.
176. Price, S.R.; Evans, P.R.; Nagai, K. *Nature* **1998**, *394*, 645.

177. Cozzone, P. J.; Jardetzky, O. *Biochemistry*, **1976**, *15*, 4853.
178. Gorenstein, D. G. In *Phosphorous-31 NMR: Principles and Applications*, Gorenstein, D. G. Ed, Academic Press, Orlando, FL, 1984, 7 – 36.
179. Moedritzer, K. *Inorg. Chem.* **1967**, *6,* 936.
180. Pietri, S.; Miollan, M.; Martel, S.; Moigne, F. L.; Blaive, B.; Culcasi, M. *J. Biol. Chem.* **2000**, *275*, 19505.
181. Jones, R. A. Y.; Katritzky, A. R. *J. Inorg. Nucl. Chem.* **1960**, *15*, 193.
182. Crutchfield, M. M.; Callis, C. F.; Irani, R. R.; Roth, G. C. *Inorg. Chem.* **1962**, 1, 831.
183. Cohn, M.; Hughes, T. R., Jr. *J. Biol. Chem.* **1960**, *237,* 3250.
184. Blumenstein, M.; Raftery, M. A. *Biochemistry*, **1972**, *11*, 1643.
185. Moon, R. B.; Richards, J. H. *J. Biol. Chem.* **1973**, *248,* 7276.
186. Porubcan, M. A.; Westler, W. M.; Ibaňez, I. B.; Markley, J. L. *Biochemistry*, **1979**, *18*, 4108.
187. van der Drift, A. C.; beck, H. C.; Dekker, W. H.; Hulst, A. G.; Wils, E. R. J. *Biochemistry*, **1985**, *24*, 6894.
188. Shan, S. O.; Herschlag, D. *Proc. Natl. Acad. Sci. USA*. **1996**, *93*, 14474.
189. Jencks, W. P. *Adv. Enzymol.* **1975**, *43*, 219.
190. Hunter, C. A.; Lawson, K. R.; Perkins, J.; Urch, C. J. *J. Chem. Soc., Perkin Trans. 2.* **2001**, 651.
191. Hunter, C. A. *Chem. Soc. Rev.* **1994**, 101.
192. Janiak, C. Dalton Trans. **2000**, 3885
193. Jennings, W. B.; Farrell, B. M.; Malone, J. F. *Acc. Chem. Res.* **2001**, *34*, 885.
194. Waters, M. *Curr. Opin. Chem. Biol.* **2002**, *6*, 736.
195. Cozzi, F.; Annuziata, R.; Benaglia, M.; Cinquini, M.; Rainmondi, L.; Baldridge, K. K.; Siegel, J. S. *Org. Biomol. Chem.* **2003**, *1*, 157.
196. Meyer, E. A.; Castellano, R. K.; Diederich, F. *Angew. Chem. Int. Ed.* **2003**, *42*, 1210.
197. Hunter, C. A. *J. Mol. Biol.* **1993**, *230*, 1025.
198. Desiraju, G. R.; Gavezzotti, A. *J. Chem. Soc. Chem. Commun.* **1989**, 621
199. Burley, S. K.; Petsko, G. A. Science, **1985**, *229*, 23.
200. Rashkin, M. J.; Waters, M. L. *J. Am. Chem. Soc.* **2002**, *124*, 1860
201. Newcomb, L. F.; Gellman, S. H. *J. Am. Chem. Soc*. **1994**, *116*, 4993.
202. Paliwal, S.; Geib, S.; Wilcox, C. S. *J. Am. Chem. Soc*. **1994**, *116*, 4497.
203. Cozzi, F.; Cinquini, M.; Annuziata, R.; Siegel, J. S. *J. Am. Chem. Soc*. **1993**, *114*, 5330.
204. Cozzi, F.; Cinquini, M.; Annuziata, D; T.; R.; Siegel, J. S. *J. Am. Chem. Soc*. **1992**, *114*, 5729.
205. Diederich, F. *Angew. Chem. Int. Ed.* **1988**, *27*, 362.
206. Leighton, P.; Cowan, J. A.; Abraham, R. J.; Sanders, J. M. K. *J. Org. Chem.* **1988**, *53*, 733.
207. West, J.; Mecozzi, S.; Dougherty, D. A. *J. Phys. Org. Chem.* **1997**, *10*, 347.
208. Kim, E.; Paliwal, S.; Wilcox, C. S. *J. Am. Chem. Soc*. **1998**, *120*, 11192.
209. Umezawa, Y.; Tsuboyama, S.; Takahashi, H.; Uzawa, J.; Nishio, M. *Tetrahedron*, **1999**, *55*, 10047.
210. Suezawa, H.; Hashimoto, T.; Tsuchinaga, K.; Yoshida, T.; Yuzuri, T.; Sakakibara, K.; Hirota, M.; Nishio, M. J. *J. Chem. Soc., Perkin Trans. 2* **2000**, 1243.
211. Doughterty, D. A.; Stauffer, D. A. *Science* **1990**, *250*, 1558.
212. Ma, J. C.; Doughterty, D. A. *Chem. Rev.* **1997**, *97*, 1303.
213. Garau, C.; Quinonero, D.; Frontera, A.; Ballester, P.; Costa, A.; Deya, P. M. *New J. Chem.* **2003**, *27*, 211.

214. Quinonero, D.; Garau, C.; Rotger, C.; Frontera, A.; Ballester, P.; Costa, A.; Deya, P. M. *Angew. Chem. Int. Ed.* **2002**, *41*, 3389.
215. Gale, P.; Navakhun, K.; Camiolo, S.; Light, M. E.; Hursthouse, M. *J. Am. Chem. Soc*. **2002**, *124*, 11228.
216. Mascal, M.; Armstrong, A.; Bartberger, M. D. *J. Am. Chem. Soc*. **2002**, *124*, 6274.
217. Kool, E. T. *Annu. Rev. Biophys. Biomol. Struc.* **2001**, 30, 1.
218. Shetty, A. S.; Zhang, J. S.; Moore, J. S. *J. Am. Chem. Soc*. **1996**, *118*, 1019.
219. Gabriel, G. L.; Iversson, B. L. *J. Am. Chem. Soc*. **2002**, *124*, 15174.
220. Ribas, J.; Cubero, E.; Luque, J.; Orozco, M. *J. Org. Chem.* **2002**, *67*, 7057.
221. Muehldorf, A. V.; Engen, D. V.; Warner, J. C.; Hamilton, A. D. *J. Am. Chem. Soc.* **1998**, *110*, 6561.
222. Hunter, C. Angew. Chem. Int. Ed. Engl. **1993**, *32*, 1584.
223. Carver, F. J.; Hunter, C. A.; Seward, E. M. *J. Chem. Soc. Chem. Commun.* **1998**, 775.
224. Adams, H.; Carver, F. J.; Hunter, C. A.; Morales, J. C.; Seward, E. M. *Angew. Chem. Int. Ed.* **1996**, *35*, 1542.
225. Hunter, C. A.; Low, C. M. R.; Rotger, C.; Vinter, J. G.; Zonta, C. *Proc. Natl. Acad. Sci.* **2002**, *99*, 4873.
226. Ishida, T.; Shibata, M.; Fuji, K.; Inoue, M. *Biochemistry* **1983**, *22*, 3571.
227. Breault, G. A.; Hunter, C. A.; Mayers, P. C. *J. Am. Chem. Soc*. **1998**, *120*, 3402.
228. Cockroft, S. L.; Hunter, C. A.; Lawson, K. R.; Perkins, J.; Urch, C. J. *J. Am. Chem. Soc*. **2005**, *127*, 8594.
229. Nicolas, L.; Beugelmans-Verrier, M.; Guilhem, J. *Tetrahedron* **1981**, *37*, 3847.
230. Stauffer, D. A.; Barrans, R. E. Jr.; Doughterty, D. A. *Angew. Chem. Int. Ed.* **1990**, *29*, 915.
231. Berger, I.; Egli, M.; Rich, A. *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 12116.
232. Sarkhel, S.; Rich, A.; Egli, M. *J. Am. Chem. Soc*. **2003**, *125*, 8999.
233. Rooman, M.; Liévin, J.; Buisine, E.; Wintjens R. *J. Mol. Biol.* **2002**, *319*, 67.
234. Wintjens R.; Liévin, J.; Rooman, M.; Buisine, E.; *J. Mol. Biol.* **2000**, *302*, 395.
235. Biot, C.; Buisine, E.; Kwasigroch, J-M.; Wintjens, R.; Rooman, M. *J. Biol. Chem.* **2002**, *277*, 40816.
236. McFail-Isom, L.; Shui, X.; Williams, L. D. *Biochemistry* **1998**, *37*, 17105.
237. Umezawa, Y.; Nishio, M. *Nucleic Acids Res*. **2002**, *30*, 2183
238. Hsu, P.; Hodel, M. R.; Thomas, W. J.; Talyor, L. J.; Hagedorn, C. H.; Hodel, A. E. *Biochemistry* **2000**, *39*, 13730.
239. Hu, G.; Gershon, P. D.; Hodel, A. E.; Quiocho, F. A. *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 7149.
240. Hu, G.; Oguro, A.; Li, C.; Gershon, P. D.; Quiocho, F. A. *Biochemistry* **2002**, *41*, 7677.
241. Boehr, D. D.; Farley, A. R.; Wright, G. D.; Cox, J. R. *Chem. Biol*. **2002**, *9*, 1209.
242. Chan, S. I.; Nelson, J. H. *J. Am. Chem. Soc*. **1969**, *91*, 168.
243. Altona, C. in *Structure and Conformation of Nucleic Acids and Protein- Nucleic Acid Interactios*, Ed. Sundaralingam, M. and Rao, S. T., University Park Press, Baltimore, USA, 1975, 613.
244. Lee, C.-H.; Ezra, F. S.; Kondo, N. S.; Sarma, R. H.; Danyluk, S. *Biochemistry* **1976**, *15*, 3627.
245. Olsthoorn, C.S.M., Bostelaar, L.J., de Rooij, J.F.M., van Boom, J.H. *Eur. J. Biochem.* **1981**, *115*, 309.
246. Simpkins, H.; Richards, E. G. *Biochemistry* **1967**, *6*, 2513.

247. Luo, R.; Gilson, H. S. R.; Potter, J.; Gilson, M. K. *Biophys. J.* **2001**, *80*, 140.
248. Warshaw, M. M.; Tinoco, I. Jr. *J. Mol. Biol.* **1965**, 13, 54.
249. Friedman, R. A.; Honig, B. *Biophys J.* **1995**, *69*, 1528.
250. Kondo, N. S.; Danyluk, S. S. *Biochemistry*, **1976**, *15*, 756.
251. Broom, A. D.; Schweizer, M. P.; Ts'o, P. O.P. *J. Am. Chem. Soc*. **1967**, *89*, 3612.
252. Kolodny, N. H.; Neville, A. C. *Biopolymers* **1980**, *19*, 2223.
253. Isaksson, J. PhD *Thesis, (ISBN 91-554-6293-6) Department of Bioorganic Chemistry, Uppsala University, Sweden* 2005.
254. Narlikar, G. J.; Herschlag, D. *Ann. Rev. Biochem*. **1997**, *66*, 19.
255. Connell, G. J.; Yarus, M. Y. *Science*, **1994**, *264*, 1137.
256. Legault, P.; Pardi, A. *J. Am. Chem. Soc*. **1994**, *116*, 8390.
257. Legault, P.; Pardi, A. *J. Am. Chem. Soc*. **1997**, *119*, 6621.
258. Muth, G. W.; Ortoleva-Donnelly, L.; Strobel, S. A. *Science*, **2000**, *289*, 947.
259. Xiong, L.; Polacek, N.; Sander, P.; Böttger, E. C.; Mankin, A. *RNA*, **2001**, *7*, 1365.
260. Drohat, A. C.; Stivers, J. T. *J. Am. Chem. Soc*. **2000**, *122*, 1840.
261. Drohat, A. C.; Stivers, J. T. *Biochemistry* **2000**, *39*, 11865.
262. Perrotta, A. T.; Shih, I.; Been, M. D. *Science*, **1999**, *286*, 123.
263. Boulard, Y.; Cognet, J. A. H.; Gabarro-Arpa, J.; LeBret, M.; Sowers, L. C.; Fazakerley, G. V. *Nucl. Acids Res.* **1992**, *20*, 1933.
264. Wang, C.; Gao, H.; Gaffney, B. L.; Jones, R. A. *J. Am. Chem. Soc*. **1991**, *113*, 5486.
265. Tollinger, M.; Crowhurst, K. A.; Kay, L. E.; Forman-Kay, J. D. *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 4545.
266. Chen, H. A.; Pfuhl, M; McAlister, M. S. B.; Driscoll, P. C. *Biochemistry*, **2000**, *39*, 6814.
267. Langsetmo, K.; Fuchs, J. A.; Woodward, C. *Biochemistry*, **1991**, *30*, 7603.
268. Ibarra-Molero, B.; Loladze, V. V.; Makhatadze, G. I.; Sanchez-Ruiz, J. M. *Biochemistry* **1999**, *38*, 8138.
269. Ha, N.-C.; Kim, M.-S.; Lee, W.; Choi, K. Y.; Oh, B.-H. *J. Biol. Chem.* **2000**, *275*, 41100.
270. Schaller, W.; Robertson, A. D. *Biochemistry* **1995**, *34*, 4714.
271. Song, J.; Laskowski, M.; Qasim, M. A.; Markley, J. L. *Biochemistry* **2003**, *42*, 2847.
272. Consonni, R.; Arosio, I.; Belloni, B.; Fogolari, F.; Fusi, P.; Shehi, E.; Zetta, L. *Biochemistry* **2003** *42, 1421*.
273. Joshi, M. D.; Sidhu, G.; Neilsen, J. E.; Brayer, G. D.; Withers, S. G.; McIntosh, L. P. *Biochemistry* **2001**, *40*, 10115.
274. Petersson, E. J. Choi, A.; Dahan, D. S.; Lester, H. A.; Dougherty, D. A. *J. Am. Chem. Soc*. **2002**, *124*, 12662.
275. Acharya, P.; Trifonova, A.; Thibaudeau, C.; Földesi, A.; Chattopadhyaya, J. *Angew. Chem. Int. Ed.* **1999**, *38*, 3645.
276. Swaminathn, S. K.; Nambiar, A.; Guntaka, R. V. *FASEB* **1998**, *12*, 515.
277. Anderson, E. M.; Halsey, W. A.; Wuttke, D. S. *Biochemistry* **2003**, *42*, 3751.
278. Nishinaka, T.; Ito, Y.; Yokoyama, S.; Shibata, T. *Proc. Natl. Acad. Sci. USA* **1997**, *94*, 6623.
279. Bar-Ziv, R.; Libehaber, A. *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 9068.
280. Bochkareva, E.; Belegu, V.; Korolev, S.; Bochkarev, A. *EMBO J.* **2001**, 20, 612.
281. Weinfeld, M.; Soderlind, K. J. M.; Buchko, G. W. *Nucleic Acids Res.* **1993**, *21*, 621.

282. Beckingham, J. A.; Cleary, J.; Bobeck, M. Glick, G. D. *Biochemistry* **2003**, 42, 4118.
283. Antson, A. A. *Curr. Opin. Struct. Biol.* **2000**, *10*, 87.
284. Handa, N.;Nureki, O.; Kurimoto, K.; Kim, I.; Sakamoto, H.; Shimura, Y.; Muto, Y.;Yokoyama, S. *Nature* **1999**, *398*, 579.
285. Deo, R. C.; Bonanno, J. B.;Sonenberg, N.; Burley, S. K. *Cell* **1999**, *98*, 835.
286. Bogden, C. E.; Fass,D.; Bergman, N.; Nichols, M. D.; Berger, J. M. *Mol. Cell* **1999**, 3, 487.
287. Antson, A.; Dodson, E. J.; Dodson, G.; Greaves, R. B.; Chen, X. P.;Gollnick, P. *Nature* **1999**, *401*, 235.
288. Messias, A. C.; Sattler, M. *Acc.Chem. Res***. 2004**, *37*, 279.
289. Jones, S.; Daley, D. T. A.; Luscombe, N. M.; Berman, H. M.; Thornton, J. M. *Nucleic Acid Res.* **2001**, *29*, 943.
290. Ellington, A. D.; Szostak, J. W. *Nature* **1990**, *346*, 818.
291. Wilson, D. S.; Szostak, J. W. *Annu. Rev.Biochem.* **1999**, *68*, 611.
292. Ramakrishnan, V. *Cell* **2002**, *69*, 557.
293. Beaudry, A. A.; Joyce, G. F. *Science* **1992**, *257*, 635.
294. Li, Y. F.;Breaker, R. R. *Curr. Opin. Struct. Biol.* **1999**, *9*, 315.
295. Jaschke, A.; Seelig B. *Curr. Opin. Chem. Biol*. **2000**, *4*, 257.
296. Feldman, A. R.; Sen, D. *J.Mol. Biol.* **2001**, *313*, 283.
297. Sidorov, A. V.; Grasby, J. A.; Williams, D.M. *Nucleic Acids Res.* **1993**, *32*, 1591.
298. Yuan, X.; Davydova, N.; Conte, M. R.; Curry, S.; Matthews, S. *Nucleic Acids Res.* **2002**, *30*, 456.
299. Tanner, J. J.; Komissarov, A. A.; Deutscher, S. L. *J. Mol. Biol.* **2001**, *314*, 807.
300. Ding, J.; Hayashi, M. K.; Zhang, Y.; Manche, L.; Krainer, A. R.; Xu, R. *Genes Dev.* **1999**, *13*, 1102.
301. Hudson, B. P.; Martinez-Yamount, M. A.; Dyson, H. J.; Wright, P. E. *Nat. Struct. Mol. Biol.* **2004**, *11*, 257.
302. Hogan, M.; Dattagupta, N.; Crothers, D. M. *Biochemistry*, **1979**, *18*, 280.
303. Lepre, C. A.; Lippard, S. J.; *Nucleic Acids and Molecular Biology*, Springer-Verlag, Berlin, 1990.
304. Zimmerman, H. W. *Angew. Chem. Intl. Ed. Engl.* **1986**, *25*, 115.
305. Steitz, T. A. *Quart. Rev. Biophys.* **1990**, *23*, 205.
306. Limmer, S.; Hofmann, H. P.; Ott, G.; Sprinzl, M. *Proc. Natl. Acad. Sci. USA* **1993**, *90*, 6199.
307. Ayer, D.; Yarus, M. *Science* **1986**, *231*, 393.
308. Du, Z.; Giedroc, D. P.; Hoffmann, D. W. *Biochemistry*, **1996**, *35*, 4187.
309. Zhu, J.; Wartell, R. M. *Biochemistry* **1997**, *36*, 15326.
310. Zhu, J.; Wartell, R. M. *Biochemistry* **1999**, *38*, 15986.
311. Toulmé J.-J. NATO ASI Series, Series A, Life Sc. 1985, 101 (Chromosomal Protein Gene Expression), 263.
312. Vesnaver, G.; Breslauer, K. J. *Proc. Natl. Acad. Sci. USA* **1991**, *88*, 3569.
313. Igoucheva, O.; Alexeev, V.; Yoon, K. *Gene Therapy* **2001**, *8*, 391.
314. Wadsworth, R. I. M.; White, M. F. *Nucleic Acids Res.* **2001**, *29*, 914.
315. Ren, J.; Chaires, J. B. *Biochemistry* **1999**, *38*, 16067.
316. Aalberts, D. P.; Parman, J. M.; Goddard, N. L. *Biophys. J.* **2003**, *84*, 3212.
317. Li, X; Liu, D.R. *J. Am. Chem. Soc*. **2003**, *125*, 10188.
318. Limmer, S.; Hofmann, H. P.; Ott, G.; Sprinzl, M. *Proc. Natl. Acad. Sci. USA* **1993**, *90*, 6199.
319. Ayer, D.; Yarus, M. *Science* **1986**, *231*, 393.

320. Ding, J.; Hayashi, M. K.; Zhang, Y.; Manche, L.; Krainer, A. R.; Xu, R. *Genes Dev.* **1999**, *13*, 1102.
321. Kurreck, J. *Eur. J. Biochem.* **2003**, *270*, 1628
322. Manoharan, M.*Biochem. Biophys. Acta.* **1999**, *489*, 117.
323. Freier, S. M.; Altmann, K. H. *Nucleic. Acids. Res.* **1997**, *25*, 4429.
324. Neuhaus, D.; Williamson, M. P. *The Nuclear Overhauser Effect in Structural and Conformational Analysis* VCH Publishers, New York, 1989.
325. Case, D. A.; Cheatham, T. E.; Darden, T.; Gohlke, H.; Luo, R.; Merz, K. M.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. J. J. Comput. Chem. 2005, 26, 1668.
326. Norberg, J.; Nilsson, L. *Acc. Chem. Res.* **2002**, *35*, 465.
327. Brunger, A. T.; Karplus, M. *Acc. Chem. Res.* **1991**, *24*, 54.

Acta Universitatis Upsaliensis

Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology 164

Editor: The Dean of the Faculty of Science and Technology

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

Distribution: publications.uu.se
urn:nbn:se:uu:diva-6741

ACTA
UNIVERSITATIS
UPSALIENSIS
UPPSALA
2006