Pure Appl. Chem., Vol. 77, No. 1, pp. 291–311, 2005.
DOI: 10.1351/pac200577010291

Electrostatic cross-modulation of the pseudoaromatic character in single-stranded RNA by nearest-neighbor interactions*

Parag Acharya and Jyoti Chattopadhyaya‡

Department of Bioorganic Chemistry, Box 581, Biomedical Center, Uppsala University, S-751 23 Uppsala, Sweden

Abstract: The generation of a single anionic or cationic center at an alkaline or acidic pH in a given molecule presents a unique opportunity to examine the electrostatic make-up of these molecules both at the neutral or ionic state. The generation of a single cationic center in the phenyl-nicotinamide system provided new straightforward evidence showing that the charge density of the electron-deficient pyridinium was actually enhanced by the donation of the charge from the electron-rich phenyl group (i.e., the pyridinyl became more basic by ca. 0.5 pK_a unit compared to an analogous system where phenyl was absent) owing to the electrostatic interactions between these two moieties. On the other hand, the generation of the 5'-guanylate ion in the hexameric single-strand (ss) RNA [5'-GAAAAC-3'], in comparison with the constituent trimeric, tetrameric, and pentameric-ssRNAs, has unequivocally shown how far the electrostatic cross-talk (as an interplay of Coulombic attractive or repulsive forces) in this electronically coupled system propagates through the intervening p**A**p nucleotide steps until the terminal p**C**-3' residue in comparison with the neutral counterpart. The footprint of the propagation of this electrostatic cross-talk among the neighboring nucleobases is evident by measurement of pK_as from the marker protons of ionization point (i.e., of **G**) as well as from the neighboring marker protons (i.e., of **A** or **C**) in the vicinity, as well as from the change of the chemical environment (i.e., chemical shifts) around their aromatic marker protons (δH2, δH8, δH5, and δH6) owing to a change of the stacking ⇆ destacking equilibrium as a function of pH.

INTRODUCTION

Hydrogen bonding, stacking interactions, hydration, and stereoelectronic effects [3–5] are some of the most important noncovalent forces [1,2] that dictate the self-assembly of nucleic acids into different architectures. It is, however, the stacking interaction that plays a more dominant role in the self-assembly of the single-stranded nucleic acid structures, which play important roles in many molecular recognition and interaction to give specific biological functions. The chemical nature of intra- and intermolecular stacking [1,2,6–21] and/or other aromatic interactions [22–27] involving both nonbiological (in the section "Aromatic interactions") as well as biological systems (in the section "Aromatic interactions in biological systems") have, therefore, recently become a topic of considerable fundamental interest to understand biological function as well as to design new molecules and materials.

---

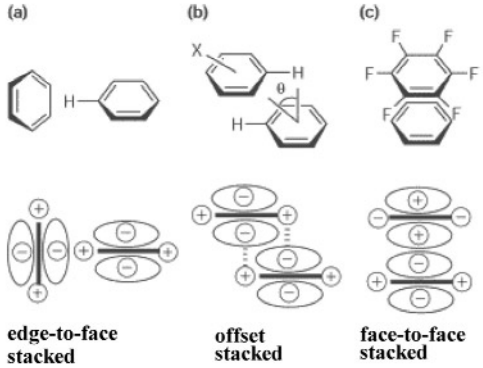
*Paper based on a presentation at the 24th International Symposium on the Chemistry of Natural Products and the 4th International Congress on Biodiversity, held jointly in Delhi, India, 26–31 January 2004. Other presentations are published in this issue, pp. 1–344.

‡Corresponding author: E-mail: jyoti@boc.uu.se

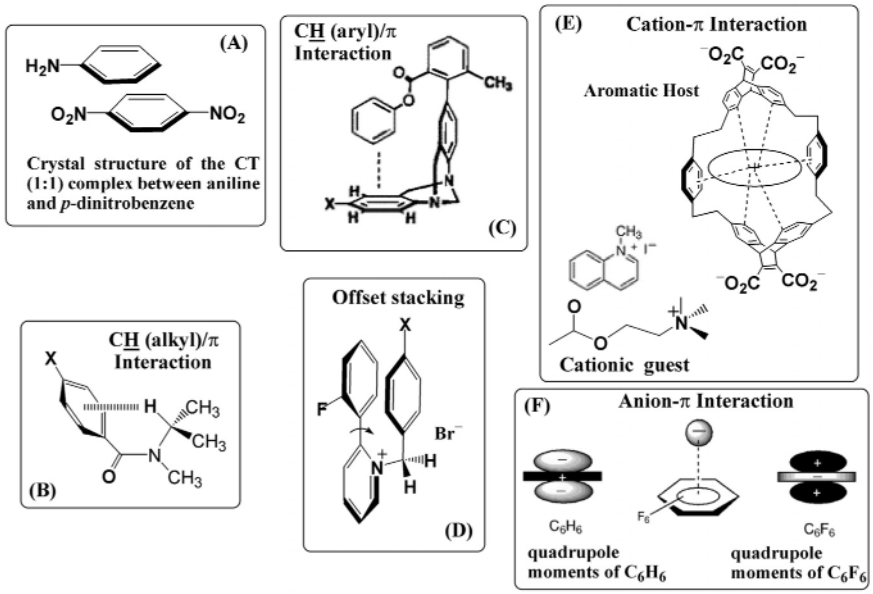
Studies using vapor pressure osmometry [15], temperature- and pH-dependent NMR [6–10,16], ORD [12], as well as theoretical simulations [11,13] have earlier been performed with nucleic acids to elucidate inter- and intramolecular base-base stacking interactions. However, these qualitative studies failed to show any insight on the electronic nature of such intramolecular base-base stacking. Some more recent studies have, however, shed considerable light on this complex problem: Recently, the dangling end stabilization studies [17–21] have estimated the strength of stacking in both DNA and RNA duplex, which has been found to be ca. 0.4–3.6 kcal mol^{-1}. According to Kool et al. [28], aromatic stacking interaction between nucleobases in water involves electrostatics (dipole–dipole and dipole-induced dipole) interactions, dispersion (momentary dipole-induced dipole) effects, and solvation. The conformational entropy opposes stacking in single-stranded nucleic acid [2] since the rotational motion across the phosphate backbone and glycosyl in a dinucleotide step is rather unconstrained. However, earlier studies qualitatively showed the sequence dependency and negligible salt effect in single-strand stacking. Very fast forward rate constant (10^7 s^{-1}) for single-strand stacking [2] and single chemical shift (NMR time-average) at any pH validates the assumption for pH-dependent two-state model (stacking $\leftrightarrows$ destacking) [9,10] in the NMR time scale.

AROMATIC INTERACTIONS

The aromatic interactions can be categorized according to the geometries [22,23,25] of interactions: face-to-face, edge-to-face (T-shaped), and offset (Scheme 1). Several studies invoked that edge-to-face and offset stacked geometries are energetically favored over the face-to-face interaction among aromatic moieties. The major noncovalent aromatic interactions (mostly in nonbiological model systems) so far identified can be grouped as follows: (i) π–π interaction [22,29], (ii) CH–π interaction (involving C<u>H</u> of both aryl [30–33] (Scheme 2C), and alkyl [34,35] (Scheme 2B), and (iii) ion–π interaction, involving both cation-π [36,37] (Scheme 2E) as well as anion-π [38–41] (Scheme 2F). Hunter et al. [22,23,42] invoked offset stacking (Scheme 1b) involving attractive atom-πσ interaction (electrostatic in nature) and edge-to-face interactions (Scheme 1a) between two aromatic moieties, rather than repulsive π–π interaction as proposed in the face-to-face stacking. However, recent studies [29] invoked an attractive face-to-face aromatic interaction (Scheme 1c) involving the negatively charged π cloud of unsubstituted benzene ring and positively charged π framework of the hexafluoro benzene (C_6F_6). Theoretical studies [43] recently showed that dispersion effects other than electrostatics dominate both aryl CH–π, and alkyl CH–π interactions. However, in all cases, alkyl CH–π interactions are weaker than aryl CH–π interactions. Dougherty et al. [37] showed that electrostatic and polarization effects are the dominant contributions in the cation–π interaction.



Scheme 1 A schematic representation of different geometries of the aromatic interactions. The negative lobe represents the π-electron cloud, and the σ-framework has been considered as positive (electron-deficient) except in the case of hexafluoro benzene [22,43].



Scheme 2 Panels A–F show the different types of aromatic interactions in nonbiological model system.

However, the quadrupole moment [37] and dispersion effect of aromatic system as well as charge-transfer interaction occasionally play a secondary role in such processes. Nishio et al. [35] proposed partial charge transfer arising from through-space proximity between alkyl hydrogen and aromatic moiety as the basis for CH–π interaction (Scheme 2B).

On the other hand, Siegel et al. [26,32,33] proposed a through-space polar (Coulombic)/π contribution as a dominating factor in the electrostatic interactions involved in edge-to-face as well as the center-to-edge (i.e., offset) oriented aromatic moieties in the neutral 1,8-diarylnapthalele system. Diedrich et al. [44] also invoked similar through-space polar/π contribution in the edge-to-face aromatic interactions. Moreover, such polar/π contribution has also been showed in aromatic interactions in the ionic states involving carboxylate ion/arene [45] and trimethylammonium ion/arene [36,46]. However, Inoue et al. [47] cited examples of ground-state partial charge transfer process in intra- and intermolecular stacking involving indole and adeninium rings. *Thus, the molecular nature of aromatic interactions still remains the subject of a major debate*.

Aromatic interactions in nonbiological model systems

Crystal structure of the CT (1:1) complex between aniline and *p*-dinitrobenzene was the first structural report [48] showing aromatic rings in a stacked arrangement (Scheme 2A). Stoddart et al. [49] first identified the edge-to-face interaction in the collapsed empty cavities of crown ethers from solid-state structure. The direct experimental evidences of intramolecular stacking in solution state have come from temperature-dependent NMR studies of side-chain-substituted dibenzodiazocine derivatives [31], concentration-dependent ^{1}H NMR studies [50] of bis-adenine with aliphatic linker, *syn*/*anti* epimerization by the temperature-dependent NMR of the 1,8-diarylnapthalene [32,33], and the dynamic NMR studies [51] of substituted benzyl pyridinium bromide. In all of these studies, the rotational free energies of model aromatic systems have been quantified by the NMR, in order to show that the edge-to-face aromatic interaction and CH–π interaction are the driving force for the observed conformational isomerism. Further, the ^{1}H NMR and solid-state studies [52] of molecular zipper complex and metal tris-bipyridine complex, as well as evaluation of the NH–π interaction-driven intermolecular association by NMR showed [53] the influence of such aromatic interaction in bimolecular complex formation.

The geometrical dependence on the strength of aromatic interaction has been observed [54] in flavoenzyme mimic. Recent studies [39] invoked a weak noncovalent attractive anion–π interaction (Scheme 2F) involving the anion and positively charged π framework of the hexafluoro benzene (C_6F_6). However, other studies [40] showed an anion–arene interaction having both positive as well as negative components. The role of heteroatoms (i.e., polar component) and solvent effect in aromatic interaction, which have been studied [50] by Gellman et al., showed the predominance of polarizibility of the aromatic surface over hydrophobicity as major attractive components in such interactions. On the other hand, Moore et al. studied [55] aromatic interaction between *m*-phenylene ethynylene-based cyclic and linear system showing the influence of hydrophobicity in such association. The solvent-dependent aromatic interactions in the foldamers [56] have been extensively studied. The NMR studies with chemical double mutant cycle system showed [23,57,58] the predominating electronic effect over dispersion forces in edge-to-face aromatic interaction, which is, however, in contrast to what Wilcox et al. observed [31].

TANDEM NEAREST-NEIGHBOR AROMATIC INTERACTIONS IN NICOTINAMIDE DERIVATIVE AS A MODEL

Yamada et al. has demonstrated [59] that by selective shielding of one side of the pyridinium face by the intramolecular face-to-face stacking (through X-ray crystallography studies) of the neighboring phenyl ring in a nicotinamide derivative (methyl derivative of **1a**), nucleophiles attack only from the nonshielded side to give exclusively 1,4-adduct over the 1,6-adduct in 99 % ee.

Electrostatic interaction-mediated pK_a perturbation of the pyridinyl group by the phenyl moiety in the phenyl-nicotinamide system

The pH-dependent ^{1}H NMR studies [60] showed that the basicity of the pyridinyl group (pK_a ~2.9–3.0) in **1a** → **1a**$^+$ (Fig. 1) could be measured both from the pH-dependent ($1 \le pH \le 7$) chemical shifts of the pyridinyl protons as well as from the protons of the neighboring phenyl and methyl groups. Thus, the pyridinium moiety in **1a** (pK_a ~2.9–3.0) becomes more basic compared to that in the standard **2a** (pK_a 2.56) due to this nearest-neighbor electrostatic interaction between phenyl and pyridinium moieties. *This observation directly provides the first straightforward chemical proof that the nearest-neighbor stacking interaction between two neighboring aromatic residues indeed cross-modulates each other's aromaticity.*

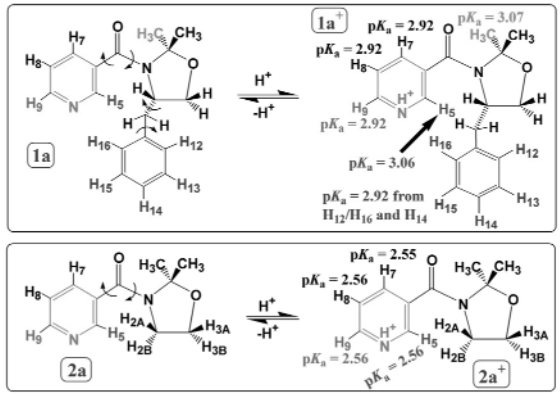


Fig. 1 Compounds used in NMR and ab initio studies [60]. The small arrow indicates possible rotational torsions. The pK_a values of pyridine moiety as calculated from each marker proton are shown.

Nearest-neighbor interaction between pyridinyl and phenyl groups

The differences in relative shielding of pyridinyl protons in protonated **1a**$^+$ compared to those in **2a**$^+$, with reference to their neutral counterparts **1a** and **2a** respectively, demonstrates that the H5/H9-edge of the pyridinyl group is more affected than the H7/H8-edge (Fig. 2) from the neighboring phenyl ring, thereby showing a direct evidence of the edge-to-face cation (pyridinium)–π (phenyl) interaction.

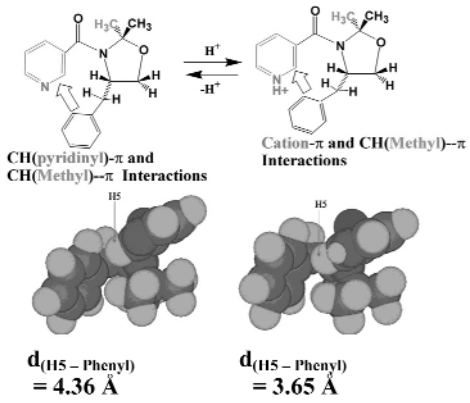


Fig. 2 Molecular modeling based on ab initio (HF/6-31G**) geometry optimization. The distance between H5 and neighboring phenyl ring $d_{(H5 - Phenyl)}$ is shown for both N- and P-states.

The quicker relaxation of pyridinyl (pyridinium) protons in **1a** or **1a**$^+$ compared to that of the standards **2a** and **2a**$^+$ shows that former relaxes through the protons of the neighboring phenyl group, respectively. The ΔT_1 (Fig. 3) shows that, for both **1a** and **1a**$^+$, the H5 relaxes more quickly than any other protons of the pyridinyl ring [*ΔpK_a ~ 0.5 from δH5, which is slightly (~0.1) higher than all other pyridinyl protons*] followed by H9 and the slowest relaxing protons being H7 and H8. This supports the preferential interaction of H5/H9 edge of pyridinyl group with the π face of the phenyl moiety. Thus, this T_1 relaxation study clearly backs up the proposed CH (pyridinyl)–π (phenyl) interaction at the N-state, and cation (pyridinium)–π (phenyl) interaction in the P-state.

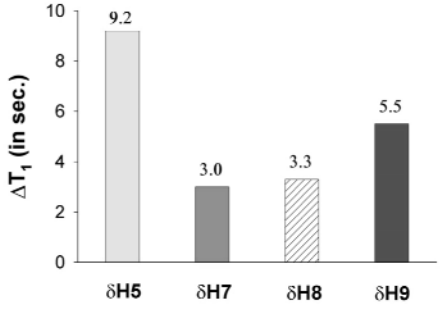


Fig. 3 The plot of ΔT_1 (where $\Delta T_1 = [T_1]_{\mathbf{1a}} - [T_1]_{\mathbf{2a}}$, in sec) at the neutral (N) state for pyridinyl protons (δH5/H9 relaxes quickly compared to δH7/δH8). The similar trend has also been found for the protonated (P) state.

The ab initio geometry optimization-based molecular modeling also supported the edge-to-face interaction in that the distance between phenyl- and H5 proton of pyridine/pyridinium groups of **1a** and **1a**$^+$ decreases by 0.7 Å upon protonation [60], thereby substantiating our observation that an attractive

electrostatic pyridinium (cation)-phenyl (π) interaction is relatively stronger (*relatively more upfield shift of $\Delta\delta H^P$ compared to $\Delta\delta H^N$, Fig. 4*) than the neutral CH (pyridinyl)–π (phenyl) interaction.

Cascade of pyridinium-phenyl-methyl cross-talk

The fact that the pH-dependent chemical shifts of the phenyl and methyl protons give [60] the pK_a of the pyridine moiety of **1a** (but not the methyl protons of **2a**) also suggests that the nearest-neighbor cation (pyridinium)–π (phenyl) interaction also steers the CH (methyl)–π (phenyl) interaction in tandem. It therefore shows that the whole pyridine-phenyl-methyl system in **1a** is electronically coupled at the ground state, cross-modulating the physicochemical property of the next neighbor. This cation (pyridinium)–π (phenyl) interaction in the P-state [60] is indeed more stable (–2.1 kJ mol^{-1}) than that of the CH (methyl)–π (phenyl) interaction (–0.8 kJ mol^{-1}) in the N-state (corroborating the chemical shift argument for $\Delta\delta H^P$ vs. $\Delta\delta H^N$ in Fig. 4).

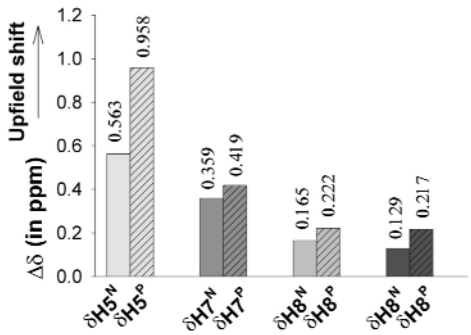


Fig. 4 The relative upfield shift (in ppm) of the pyridinyl protons (H5/H7/H8/H9) in presence of neighboring phenyl in **1a** compared to its absence in **2a** at the N-state ($\Delta\delta H^N > 0$) or at the P-state ($\Delta\delta H^P > 0$) where $\Delta\delta H^N = \delta H_{\mathbf{2a}} - \delta H_{\mathbf{1a}}$ and $\Delta\delta H^P = \delta H_{\mathbf{2a}}\mathbf{+} - \delta H_{\mathbf{1a}}\mathbf{+}$.

IMPLICATIONS

The present pH-dependent ^{1}H NMR study has given straightforward experimental evidence of the intramolecular cation–π interaction in edge-to-face geometry, thereby perturbing the pK_a of the pyridine-nitrogen in the presence of the neighboring aromatic group. Thus, this can be used as a major experimental tool to identify the nature of aromatic interaction in a system having a protonation or a deprotonation site. Interestingly, such a titration method also gives an enormous insight into the energetics, geometry as well as the nature of weak noncovalent contribution in such aromatic interactions, which are of great importance in molecular recognition pattern in both biological as well as nonbiological systems.

AROMATIC INTERACTIONS IN BIOLOGICAL SYSTEMS

Evidence regarding the nature of intramolecular aromatic interactions in nucleic acids and their complexes [42,61–80] have mainly come from various structural studies. Thus, Hunter et al. first invoked the presence of offset stacked nucleobases in DNA [42] based on the X-ray crystallographic data followed by computer modeling to construct conformation-dependent energy maps based on van der Waals and electrostatic interactions calculated between stacked bases. Recent crystallographic studies in RNA have proposed the presence of stabilizing O4' (lone pair)–π (nucleobase) [66], water (lone pair)–π (nucleobase) [67], as well as water (H)–π (nucleobase) [67] from the proximity calculations. Rooman et al. [68,69,80] defined and analyzed stair-shaped motifs, which simultaneously involve base

stacking, hydrogen bonding, and cation–π interaction in protein-DNA complex through the geometrical proximity found in the X-ray crystallographic database. Recent database studies [71] showed the importance of thymine-methyl/π interaction in the sequence-dependent deformability of DNA. Moreover, studies [70] based on screening of nucleic acid databases showed that divalent cations [like $Mg(OH_2)_6^{2+}$] interact favorably with π systems of nucleic acid bases. It has also been proposed [70] that some critical cation–π interactions may contribute to the stability of the anticodon arm of yeast $tRNA^{phe}$ and to the magnesium core of the Tetrahymena group I intron P4-P6 domain. Such cation–π interactions have been also implicated [70] in DNA bending, DNA-protein recognition, base flipping, RNA folding, and catalysis. Ab initio studies have shown the presence of the aromatic interactions (mostly of cation–π in nature) [77] between protein and DNA involving positively charged Arg or Lys side chains and aromatic rings of nucleic acids. The X-ray studies along with calorimetric and fluorescence analyses have shown the importance of electrostatic cation–π interaction in the protein recognition of the m^7G part of the mRNA cap structure [63–65]. Similar kinetic and calorimetric experiments [73] have also identified the key aromatic π–π stacking interaction between Tyr41 and the adenine ring of bound nucleotides in the active site of an aminoglycoside phosphotransferase enzyme. The importance of edge-to-face aromatic interactions in preorganization of the peptide secondary structure has been extensively studied by Waters et al. [74–76].

pK_a PERTURBATION IN NUCLEIC ACIDS AND PROTEIN FOLDING

pK_a perturbation in nucleic acids

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

Structural studies [82–95] have revealed that in a large number of RNAs (and in certain cases in DNAs), the pK_a values of nucleobases (particularly adenine and cytosine) are significantly perturbed relative to that of mononucleotides. Thus, the most effective acid-base catalysis can be performed with ionizable groups having pK_a near the physiological pH of 7, thereby accounting for widespread use of histidines in protein enzyme active sites [81]. NMR studies with leadzyme showed the perturbed pK_a of 6.5 for the adenine (A25) at the active site [83,84]. Similarly, adenines within the internal loop at domain B of the hairpin ribozyme showed perturbed pK_a values ranging from 4.8 to 5.8 [85]. The shifted adenine pK_a value 6.2 for A^+.C wobble base pair have been found in an internal loop in domain A of the hairpin ribozyme [86]. A perturbed pK_a ~7.6 for an adenine (A2451) in the ribosomal peptidyl transferase center has been reported [87]. However, research by Xiong et al. [88] showed contradictory results, which proposed pH-independent behavior of A2451 dimethylsulfate modification in ribozyme of *Thermus aquaticus* and *Mycobacterium smegmatis*. The perturbed pK_a 6.4 for uracil in the ternary uracil DNA glycosylase complex has also been reported [89,90]. Studies from Nakano et al. [91] as well as Lupták et al. [92] showed moderate perturbation of pK_a to near neutrality for active-site cytosine (C75) in HDV ribozyme. Similar pK_a perturbation for cytosine has also been observed [93] from studies on self-cleaving ribozyme. Similarly, The A^+.C wobble base pairing involving adenine with perturbed pK_a 6.6 has also been reported [94,95] for DNA oligomer.

The shift of the pK_a values of nucleobases at the active site in ribozyme toward neutral pH (in ground- rather than transition-state perturbation [96,97]) might arise from: (i) the stabilization by H-bonding (by wobble base pairing [83–85,94,95]), (ii) salt-bridge formation [84] or stabilizing (for the protonated nucleobase) and destabilizing (for the deprotonated nucleobase) interaction with the phosphodiester moiety in the steric proximity, and (iii) the change in the structural microenvironment (by folding or unfolding) [97].

Thus, it is likely that a local hydrophobic pocket around a specific nucleobase (created by folding) may reduce the local dielectric of the microenvironment [97,98], thereby increasing the pK_a values. Alternatively, a relatively greater exposure of a specific nucleobase to the aqueous medium may give a pK_a value very similar to that of a mononucleotide counterpart. This is consistent with the fact [98] that the pK_a in general decreases (i.e., more acidic) as the dielectric of the medium increases. A particular example of this phenomenon can be found [98] in the recent determination of pK_a for –COOH group in salicycilc acid, which showed pK_a of 6.6 in DMSO ($\varepsilon = 48$) and 2.9 in water ($\varepsilon = 78$).

However, the proton is sequestered by the H-bonding interaction through which its pK_a perturbation occurs (i.e., by i and ii, see above), thereby becomes ineffective in catalytic participation. On the other hand, the pK_a perturbation via local microenvironment change (as described in iii) is considered to be a more efficient way to steer the enzymatic properties (both in general acid-base catalysis [96,97,99,100] and electrostatic catalysis [96,81]) of RNA.

It has been commented [97] that limited folding pattern (so far discovered), lesser diversity of polar side chains, and negatively charged (hydrophilic) backbone in RNA compared to those of protein, apparently, restrict the ability of the former to modulate the microenvironment at the binding sites. On the contrary, our pH-dependent NMR studies with ssRNAs (see the section "Cross-modulation of physicochemical character of nucleobases in single stranded ribonucleic acids"), for the first time, has experimentally demonstated [136–138] that the microenvironment around the nucleobases of RNA varies widely, and consequently, their intrinsic pseudoaromatic characteristics change in a variable manner depending upon their differential electrostatic modulation through sequence-dependent nearest-neighbor offset stacking interaction, which forms the current basis of the concept of the extended genetic code (see the section "Implications").

pK_a perturbation in proteins

Unusual pK_a values have been observed for many protein residues [97,101–111] at the (i) active site and implicated in catalytic function, (ii) ligand-binding site, and (iii) protein–protein interaction site. Such pK_a perturbation may arise due to changing electrostatic properties of the binding surface, which modulates (a) charge–charge interaction between ionizable groups, (b) differential solvent accessibility of charged side chains, and (c) H-bonding interactions. For example, a pK_a 6.36 for glutamate side chain in rat CD2 has been attributed [103] to mutual electrostatic repulsion. Aspartate side chain pK_a of 7.5 in oxidized *Escherichia coli* thioredoxin has been explained [104] in terms of hydrophobic local environment. A perturbed pK_a at the binding site of nicotinic acetylcholine receptor [111] has recently been shown. The pK_a of histidine residues at the C-terminus and N-terminus of the folded α-helices of barnase protein are shown [101] to be ~0.5 units higher and 0.8 units lower, respectively, compared to the unfolded residue. This perturbation of pK_a results from the combination of the different electrostatic environments. The increase of pK_a at C-terminus is due to the charge-helix dipole interaction and stabilization through H-bonding, whereas the lowering of pK_a has been attributed to the movement of the side chain away from the protein, thereby promoting solvent-induced electrostatic screening. The positioning of a charged group affects its pK_a in protein as shown [97] in the active-site Lys of acetoacetate decarboxylase, in which its pK_a is perturbed by ~4 units to a value of 6 due to the presence of a nearby Lys residue.

Recent NMR titration studies [107,108] with turkey ovomucoid third domain (OMTKY3) and DNA binding proteins from archaeum *Sulfolobus solfataricus* (Sso7d) [109] showed that pK_a of a particular residue was sensed by resonances not only in that residue, but also from neighboring residue. In OMTKY3, the major interactions [108] responsible for lowering of pK_a of $P_{1'}$-Glu19 are H-bonding between carboxylate of $P_{1'}$-Glu19 and hydroxyl group of P_2-Thr17, the charge–charge interaction between $P_{1'}$-Glu19 and P_3-Arg21, and the intra-residue H-bonding between carboxylate of $P_{1'}$-Glu19 and its own amide group. However, the elevated pK_a of $P_{25'}$-Glu43 might arise due to helix–dipole interaction. The

perturbation of pK_as of the different side chains in Sso7d [109] are due to the H-bonding (between Glu35 and Tyr33 residues), electrostatic screening as well as formation of salt bridge (between Asp34 and Tys20 residues) within the active site. Thus, the local electrostatic-mediated pK_a perturbation has been shown [107–110] to contribute site-specifically to the pH dependence of the protein stability. Several other studies [102–106] have come up showing the importance of the electrostatic interaction of neighboring residues in the overall stability and functions of proteins.

BIOLOGICAL IMPORTANCE OF SINGLE-STRANDED NUCLEIC ACIDS

The single dangling nucleotide [17–20] at the end of both DNA and RNA duplexes is known to increase the duplex stability. In more recent studies [21], it has been shown that longer single-stranded dangling residues (up to tetranucleotide) stabilize the RNA–RNA and DNA–DNA duplexes even slightly more (by an extra ~0.1–1.0 kcal mol^{-1}) than the single-nucleotide dangling end (~2.0 kcal mol^{-1}). The ssRNA [18,21] as dangling end stabilizes the helix, which results in specific biological function: (i) The CCA overhang at the 3'-terminal of tRNA, which is involved in aminoacylation reaction, also stabilizes [112] the cloverleaf structure of tRNA. (ii) The dangling end adjacent to the codon-anticodon base pair stabilizes [113] the interaction between mRNA and tRNA. (iii) The dangling nucleotide at the 3'-end of a pseudoknot RNA is also known [114] to stabilize the stem structure. (iv) Single unpaired base bulges in RNA duplexes enhance [115,116] the stability of the RNA more compared to the fully base-paired counterpart, in which both the base identity as well as the nearest-neighbor context have been shown to be important for the overall stability of the bulges. (v) Recognition and interaction with many ligands including proteins also take place with ssRNA.

The comparison of ssDNA dangling end with that of ssRNA showed [19] that the ssDNA motif with 5'-dangling ends or ssRNA with 3'-dangling ends contribute to the stability of the duplex equally or more than their ssRNA or ssDNA counterparts, respectively. It has also been shown [20] that polarizibility of dangling base correlates with the stability of duplex better than the hydrophobicity.

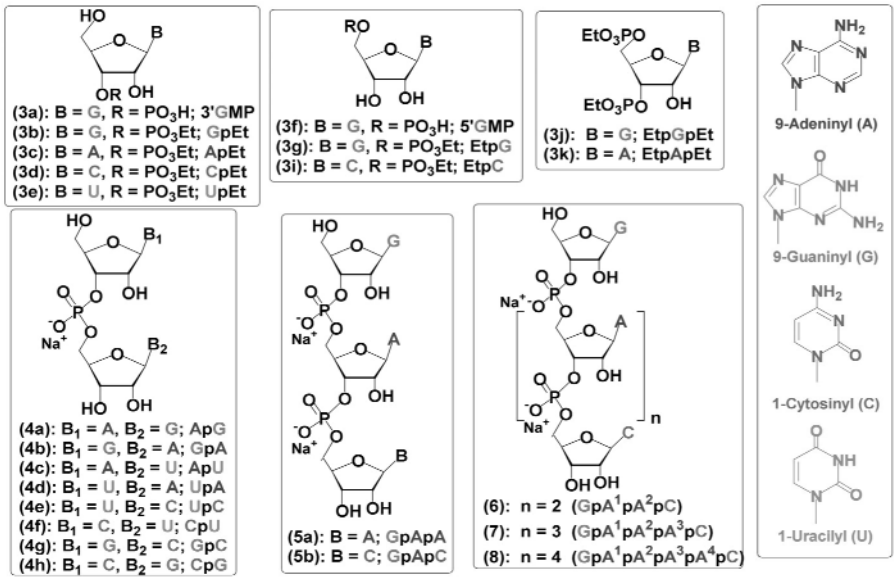
Many ssDNAs also show [117–130] their functional properties upon binding to specific proteins. These ssDNA binding proteins belong to either of the two following categories: (i) those that recognize a particular sequence of nucleic acids like transcriptional regulation [130] or telomere replication [117], and (ii) those that specifically interact with a particular physical form of nucleic acid [118,122,126] in a sequence-independent manner like *E. coli* ssDNA binding protein, RepA protein [126], or RecA protein [118,122]. The stacking interaction between aromatic side chain of protein and nucleobase is considered to be a major force in the binding process [119] for those proteins in the second category, which bind to the single-stranded nucleic acids (SSB proteins). Stacking between aromatic amino acids and nucleic acid bases (mainly enthalpic in nature) plays an important role [121] in the enzyme specificity with nucleic acid substrate. In recent studies [131–135], four main ssRNA binding proteins which act sequence-specifically have been recognized: sex-lethal protein from *D. melanogaster* [132], N-terminal domain of polyA binding protein (PABP) [133], Trp RNA-binding attenuation protein (TRAP) [134], and transcription termination factor Rho protein [135], which binds to a particular ssRNA sequence imparting specific biological functions. In all of these interactions, it has been proposed [131] that sequence-specific recognition through RNA bases determines the specificity.

CROSS-MODULATION OF PHYSICOCHEMICAL CHARACTER OF NUCLEOBASES IN SINGLE-STRANDED RIBONUCLEIC ACIDS

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

pH-dependent titration of di-, tri-, tetra-, penta-, and hexameric ssRNAs

The di-ribonucleotides **4a–f** (Scheme 3) for the pH-dependent ^{1}H NMR studies (at 298 K) are chosen [136] such that only one of the two nucleobases in the molecule can be exclusively an alteration of the electronic character of one aglycon on its nearest neighbor. The pH titration by ^{1}H NMR for each nucleobase in **4a–f** shows [136] that pK_a of a particular nucleobase can be obtained (Table 1) from not



Scheme 3

Table 1 The pK_a and related free energy of protonation (ΔG^o_{pKa}) from the ^{1}H NMR titration of aromatic marker protons for di-ribonucleotides (**4a–f**) and corresponding monomeric nucleoside 3'-ethylphosphates (**3b–e**).

Compounds		pK_a and ΔG^o_{pKa} from aromatic marker protons[a]													
		δ H8G		δH8A		δH2A		δH6C		δH6U		δH5C		δH5U	
		pK_a	ΔG^o_{pKa}	pK_a	ΔG^o_{pKa}	pK_a	ΔG^o_{pKa}	pK_a	ΔG^o_{pKa}	pK_a	ΔG^o_{pKa}	pK_a	ΔG^o_{pKa}	pK_a	ΔG^o_{pKa}
ApG (4a)	5'A^{H+}	1.64	9.0	2.88	16.4	2.83	16.2	–	–	–	–	–	–	–	–
	3'G^-	9.42	53.7	9.71	55.4	9.65	55.1	–	–	–	–	–	–	–	–
GpA (4b)	5'G^-	9.17	52.4	9.11	52.0	9.11	52.0	–	–	–	–	–	–	–	–
	3'A^{H+}	1.68	9.6	3.22	18.4	2.94	16.8	–	–	–	–	–	–	–	–
ApU (4c)	5'A^{H+}	–	–	2.95	16.8	2.95	16.8	–	–	2.98	17.0	–	–	2.95	16.8
	3'U^-	–	–	9.35	53.3	9.33	53.2	–	–	9.36	53.4	–	–	9.42	53.7
UpA (4d)	5'U^-	–	–	_[b]	_[b]	_[b]	_[b]	–	–	9.09	51.9	–	–	9.09	51.9
	3'A^{H+}	–	–	3.07	17.5	3.06	17.5	–	–	3.12	17.8	–	–	3.01	17.2
UpC (4e)	5'U^-	–	–	–	–	–	–	9.14	52.1	9.06	51.7	9.06	51.7	9.04	51.6
	3'C^{H+}	–	–	–	–	–	–	3.71	21.2	3.71	21.2	3.71	21.2	3.71	21.2
CpU (4f)	5'C^{H+}	–	–	–	–	–	–	3.56	20.3	3.48	19.9	3.58	20.4	3.58	20.4
	3'U^-	–	–	–	–	–	–	9.18	52.4	9.21	52.5	9.18	52.4	9.25	52.8
GpEt (3b)	G^-	9.25	52.8	–	–	–	–	–	–	–	–	–	–	–	
ApEt (3c)	A^{H+}	–	–	3.11	17.7	3.10	17.7	–	–	–	–	–	–	–	
CpEt (3d)	C^{H+}	–	–	–	–	–	–	3.84	21.8	–	–	3.84	21.8	–	
UpEt (3e)	U^-	–	–	–	–	–	–	–	–	9.44	53.9	–	–	9.43	

[a]pK_a calculated from respective Hill plot analyses. The error for pK_a is in between ±0.01 to ±0.04 and the corresponding error in ΔG^o_{pKa} is between ±0.1 to ±0.2 kJ mol^{-1}. ΔG^o_{pKa} = 2.303*RT* pK_a, where R = 8.314 J K^{-1} mol^{-1} and T = 298 K.
[b]No titration profile.

only its own aromatic maker protons, but also from the marker protons of the neighboring nucleobase as a result of cross-modulation of two-coupled π systems of protonated (for **A**, **G**, or **C**) or deprotonated (**G** or **U**) at a given pH in order to show the effect of neighboring nucleobases. This also shows the sequence-specific effect in two possible isomeric aglycone combinations in purine-pyrimidine (**4c** and **4d**), purine–purine (**4a** and **4b**), as well as pyrimidine–pyrimidine (**4e** and **4f**) dimers. Similarly, pH-dependent titration (over pH 7.2–12.2 at 298 K) using ^{1}H NMR has been performed for tri-, tetra-, penta-, and hexameric ssRNAs (**5a**, **5b**, **6**–**8**, Scheme 3) for which the sequences are chosen in such a way that over the whole titration range, only single-base ionization effect (at 5'-**G**) can be observed across the strand.

Sequence-specific nearest-neighbor interaction and thermodynamics of the offset stacking

The pH-dependent chemical shift change [$\Delta\delta_{(N-D)}$, Fig. 6] in any of the aromatic marker protons in either of the two coupled nucleobases in dinucleotide monophosphates (**4a**–**f**) shows *variable* electrostatic modulation depending upon the geometry of the offset stacking, partial charge of the heteroatom as well as the sequence. The thermodynamics of the pH-dependent offset stacking in di-ribonucleotides (**4a**–**f**) are shown [136] in Fig. 5. Among all these dimers, in CpU$^-$ the formation of 1-uracilylate as a function of pH failed to promote any destacking ($\Delta\delta_{N-D} > 0$, see Fig. 5 and [$\Delta G^o{}_{stacking}]_{N\text{-}state}$: –2.3 kJ mol^{-1} compared to [$\Delta G^o{}_{stacking}]_{D\text{-}state}$: –1.9 kJ mol^{-1}). In general, the extent of stacking in di-ribonucleotides decreases (Fig. 5) [136] in the following order: 5'-purine–purine-3' ≅ 5'-purine–pyrimidine-3' > 5'-pyrimidine–pyrimidine-3' > 5'-pyrimidine–purine-3' at the N-state. Similarly, the pairwise comparison of the $\Delta\delta_{(N-D)}$ (Fig. 7) of all aromatic marker protons [137,138] of tri-, tetra-, penta-, and hexameric ssRNAs (**5a**, **5b**, **6**–**8**, Scheme 3) shows how the edge of the nucleobase responds to the interplay of the nearest-neighbor electrostatic interactions due to 9-guanylate formation. The δH8A in ssRNA is relatively more destacked [$\Delta\delta_{(N-D)} < 0$, Fig. 7] [138], showing its significant pH-dependent change (i.e., greater modulation), compared to that of δH2A. We have argued on the basis of this variable $\Delta\delta_{(N-D)}$ (Fig. 7) that the increase of pK_a of 9-guanylate observed from the neighboring nucleobases is owing to their enhanced electron density, which is caused by electrostatic interactions [137,138] between two coupled nucleobases. That the electrostatic interactions between two neighboring aromatic systems (one is electron-rich and the other is electron-deficient) involve an actual charge transfer is also evident from the stacked pyridine–phenyl interactions in the nicotinamide derivative **1a** (see the section "Electrostatic interaction-mediated pK_a perturbation of the pyridinyl group by the phenyl moiety in the phenyl-nicotinamide system") [60], in which the pyridinyl group is relatively electron-deficient,

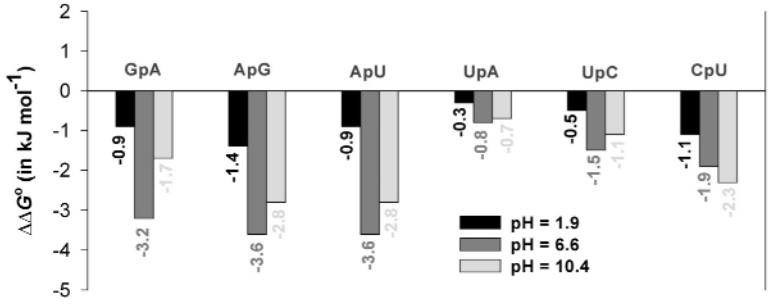


Fig. 5 The free energy of stacking ($\Delta\Delta G^o \cong \Delta G^o_{stacking}$, in kJ mol^{-1}) [137] for **4a**–**f** at neutral (pH = 6.6), protonated (pH = 1.9), and deprotonated (pH = 10.4) state. $\Delta G^o_{stacking}$ has been estimated using the relation: $\Delta\Delta G^o \cong \Delta G^o_{stacking} = [\Delta G^o_{N/S(298\ K)}]_{\mathbf{Np}N'} - [\Delta G^o_{N/S(298\ K)}]_{\mathbf{Np}Et}$ where **Np**N' and **Np**Et signify 5'-nucleotidyl moiety (in bold) of di-ribonucleoside monophosphate (**Gp**A, **Ap**G, **Ap**U, **Up**A, **Up**C, and **Cp**U; see Scheme 3) and nucleoside 3'-ethylphosphate (**Np**Et with N = A, G, U, and C; see Scheme 3), respectively.

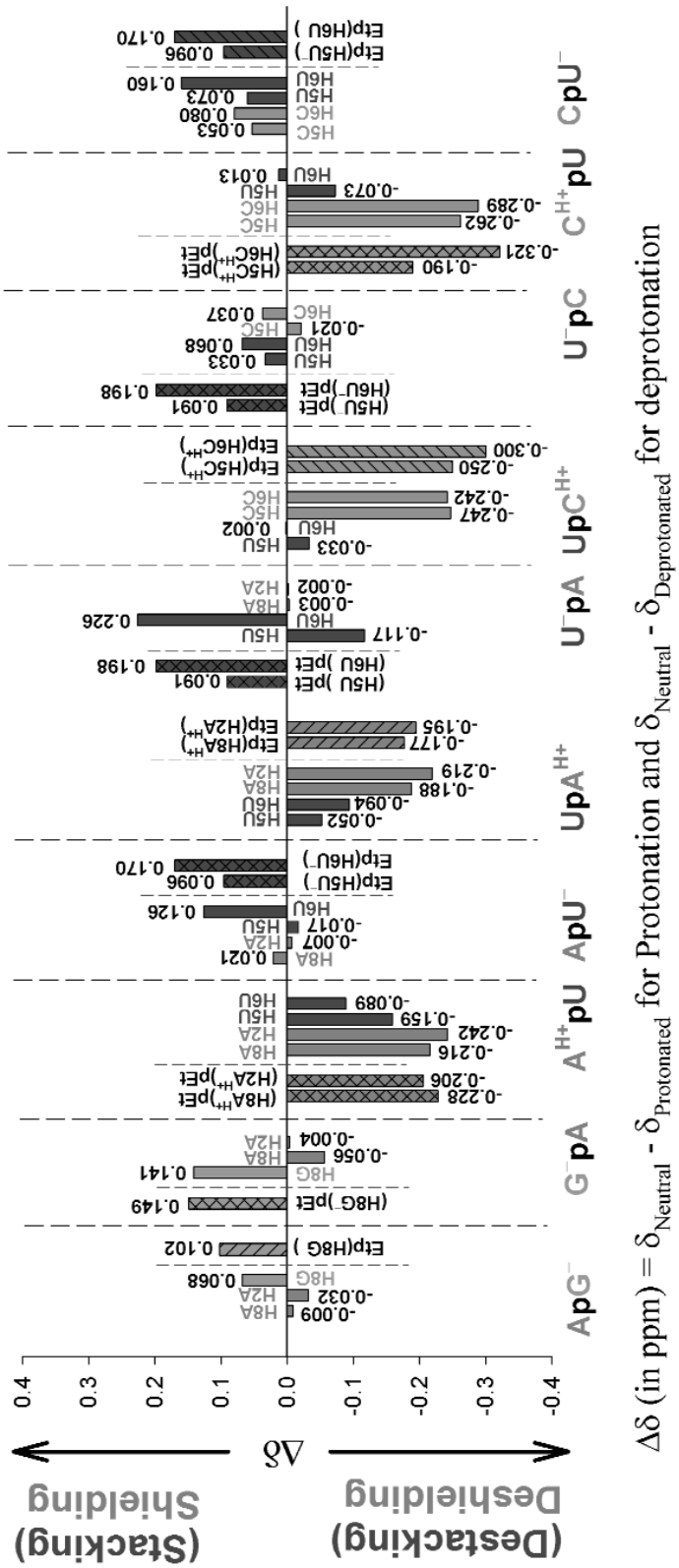


Fig. 6 Sequence-specific nearest-neighbor interaction in dinucleotides.

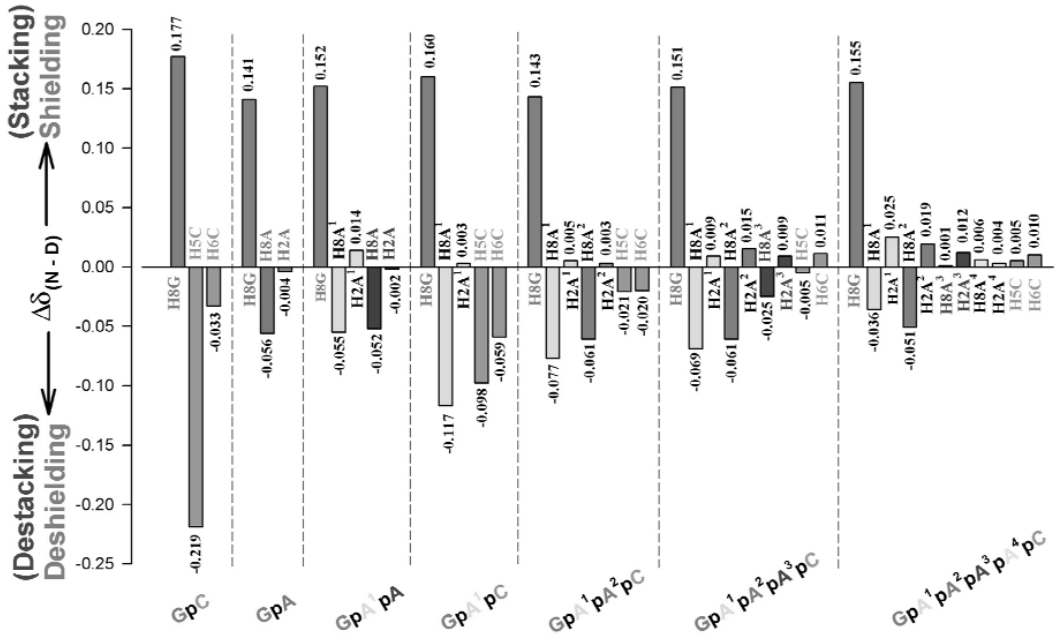


Fig. 7 Effect of the deprotonation of N^1H guanin9-yl.

whereas the phenyl group is relatively electron-rich. Consequently, the pK_a of the pyridinyl moiety in **1a** was found to increase by ~0.5 pK_a unit because of electrostatic interactions with the neighboring phenyl group, compared to the analog **2a** in which phenyl was absent. The electron-donating character of the phenyl group in the stacked pyridine–phenyl interactions in the nicotinamide derivative **1a** is also evident by the relative downfield shift of the phenyl group in **1a**$^+$ compared to **1a**. The absence of any characteristic charge-transfer band in the UV absorption in neither of the single-stranded dimers or oligomers or in nicotinamide derivative suggests that the potential difference between the donor and acceptor in this electrostatic charge transfer must be very low.

pK_a shift of 9-guanylate due to electrostatic effect of 3'- and 5'-phosphate vs. nearest-neighbor nucleobase

Comparison of the pK_a of 9-guaninyl in GpEt (**3b**) and EtpG (**3g**) shows (Table 2) that the 5'-phosphate in the former makes the pK_a of 9-guaninyl more basic ($[\Delta pK_a]_{(\mathbf{3g})-(\mathbf{3b})} \approx 0.32$) compared to the 3'-phosphate in the latter. This is due to the electrostatic interaction (repulsion) as a result of the spatial proximity of the negatively charged 5'-phosphate and the fused imidazole moiety of the 9-guaninyl system (in the *anti* conformation), which causes an effective decrease of the electron density in the imidazole part [δ_{H8} = 8.078 ppm in EtpG (**3g**) compared to 8.01 ppm in GpEt (**3b**)] by donating the charge in the pyrimidine moiety, thereby producing an overall increase of pK_a of the N^1 of the 9-guaninyl by 0.32 unit. The pairwise comparison of pK_a of 9-guanylate (**G**$^-$) at the dimer and trimer level, however, shows (Tables 1 and 2): in **G**pA (**4b**)/**G**pC (**4g**) giving $[\Delta pK_a]_{(\mathbf{4g})-(\mathbf{4b})} \approx 0.39$; similarly, in **G**pApA (**5a**)/**Gp**ApC (**5b**) having same phosphate charge shows $[\Delta pK_a]_{(\mathbf{5a})-(\mathbf{5b})} \approx 0.13$. However, in Ap**G** (**4a**)/Cp**G** (**4h**) giving $[\Delta pK_a]_{(\mathbf{4h})-(\mathbf{4a})} \approx 0.03$. Thus, the p$K_a$ of N^1-H of 9-guaninyl residues is sequence-dependent since the two isomeric dimers and/or trimers have the same phosphate charge, but a different 3'-nucleobase. This suggests that the chemical nature of the nucleobase modulates the physicochemical make-up of the nearest-neighbor nucleobase more effectively than the phosphates, although the latter being negatively charged also imparts electrostatic effect through phosphate–nucleobase inter-

action (depending upon conformation across γ). Moreover, $[\Delta pK_a]_{(\mathbf{4g})-(\mathbf{4b})} \approx 0.39$ compared to $[\Delta pK_a]_{(\mathbf{4h})-(\mathbf{4a})} \approx 0.03$ shows that nearest-neighbor modulation in p**G** is opposed by the 5'-phosphate effect when compared to **G**p (with 3'-phosphate) in similar dimeric sequences.

Table 2 The pK_a and ΔG^o_{pKa} of 9-guanylate ion (**G**) from the aromatic marker protons in di- and tri-ribonucleotides (**4g**, **4h**, **5a**, and **5b**) and their corresponding monomeric compounds (**3a**, **3b**, **3f**, **3g**, and **3j**) showing the electrostatic effect of 3'- and 5'-phosphate vs. nearest-neighbor nucleobase.

Compounds		pK_a and ΔG^o_{pKa} of the 9-guanylate ion (G)[a,b]			
		δH8		δH2 / δH5 / δH6	
		pK_a	ΔG^o_{pKa}	pK_a	ΔG^o_{pKa}
GpEt (**3b**)		**9.25**	**52.8**	–	–
3'-**G**MP (**3a**)		**9.33**	**53.2**	–	–
5'-**G**MP (**3f**)		**9.74**	**55.6**	–	–
Etp**G**pEt (**3j**)		**9.57**	**54.6**	–	–
5'-Etp**G** (**3g**)		**9.57**	**54.6**	–	–
GpC	**G**p	9.56	**54.5**	–	–
(**4g**)	pC	–	–	**9.56**[d]	**54.5**
				9.51[e]	**54.3**
Cp**G**	Cp	–	–	**9.42**[d]	**53.7**
(**4h**)				**9.42**[e]	**53.7**
	p**G**	**9.45**	**53.9**	–	–
GpApA	**G**p	**9.75**	**55.6**	–	–
(**5a**)	pAp	**9.74**	**55.6**	**9.82**	**56.0**
	pA	**9.78**	**55.8**	–[c]	–[c]
GpApC	**G**p	**9.88**	**56.4**	–	–
(**5b**)	pAp	**9.88**	**56.4**	–[c]	–[c]
	pC			**9.89**[d]	**56.4**
				9.90[e]	**56.5**

[a]pK_as calculated from Hill plots[ref]; average error ±0.01 to 0.02.
[b]ΔG_{pKa} (in kJ mol^{-1}) = 2.303*RT*pK_a ; average error ±0.1.
[c]No titration profile.
[d]For δH5.
[e]For δH6.

Variation of pK_a of 9-guaninyl among different marker protons across the single strand

Intrinsic pK_a (Fig. 8) at a specific site is the one arising directly from a single ionizable group without any electrostatic interference of any other titrable or electronic group, as observed in the pK_a of N^1 of **G** in the monomeric **G**pEt (**3b**). However, the presence of electrostatic interaction of **G** with the other electronic groups such as neighboring nucleobases, phosphates, and the pentose-sugar units across the single strand causes the modulation/perturbation of the intrinsic pK_a of the **G** (termed as apparent pK_a–1, where $\Delta pK_a = (pK_a \text{ of } \mathbf{G})_{\text{from δH8G in ssRNA}} - (pK_a \text{ of } \mathbf{G})_{\text{from δH8G in GpEt}}$, Fig. 8), as found in the di-, tri-, tetra-, penta-, and hexameric ssRNAs (**4a**, **4b**, **5a**, **5b**, **6**–**8**, Scheme 3), which can be quantified by comparison with that of the respective monomer **G**pEt (**3b**). In oligo-ssRNAs, the marker proton from neighboring nucleobases (**A** or **C**), as well as the ionic phosphates (which are nontitrable over

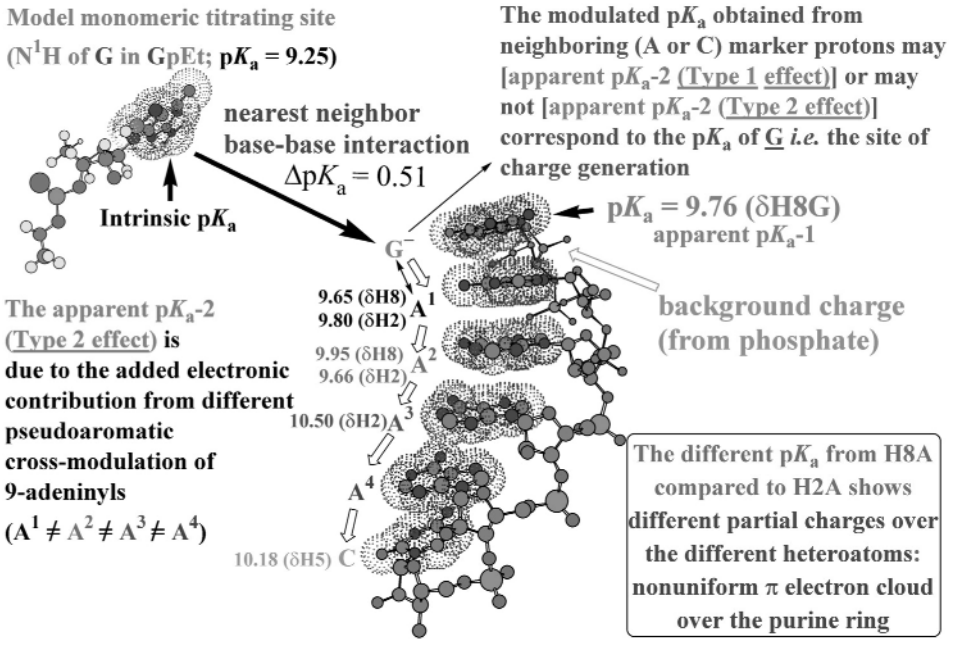


Fig. 8 A schematic representation of the nearest-neighbor electrostatic interaction in hexameric ssRNA (**8**), thereby showing the pK_a perturbation due to differential modulation of the local microenvironment across the single stand.

the pH range 7.2–12.5) also show the pK_a of **G** [termed as apparent pK_a–2 where ΔpK_a = (pK_a of **G**)$_{\text{from } \delta H8G \text{ in ssRNA}}$ − (pK_a of **G**)$_{\text{from marker protons (H8A/H2A/H5C/H6C) of ssRNA}}$, Fig. 8] because of the variable electrostatic interaction of **G** /**G**$^-$ with those neighboring electronic groups. Three distinct cases of variation of apparent pK_a–2 (with respect to apparent pK_a–1) have been observed in this work: (i) when apparent pK_a–2 > apparent pK_a–1, it implies an additional electrostatic energy input from the electronic character of the neighboring nucleobase (Type 2 effect) [137,138]; (ii) when apparent pK_a–2 < apparent pK_a–1, it suggests an electrostatic screening by the solvent; (iii) when apparent pK_a–1 = apparent pK_a–2, it means no additional electrostatic energy input from the electronic character of the neighboring nucleobase (Type 1 effect) [137,138]. Thus, the apparent pK$_a$–2 (Type 2 effect) [137,138] of the **G** gives the strength of its cross-modulation through the differential electrostatic interaction with the variable electronic nature (pseudoaromaticity) of the neighboring nucleobases. Both the apparent pK_a–1 and apparent pK_a–2 are site-specific, and thus microscopic in nature [138].

The pK_a of 5'-**G**p (apparent pK_a–1) residue in oligo-ssRNAs as observed from its own δH8G varies from 9.76 ± 0.01 (in **6**) to 9.88 ± 0.01 (in **8**). The variation of pK_as [ΔpK_a] for 5'-**G**p residue as measured from the other aromatic marker protons [apparent pK_a–2 (Type 2 effect)] of various nucleotide residues [138] across the strand in ssRNA: $\Delta pK_a \approx 0.9$ and 0.14 for hexameric and pentameric ssRNAs (**8** and **7**), respectively. Similar ΔpK_as [apparent pK_a–2 (Type 2 effect)] for tri- and tetrameric ssRNA (**5** and **6**) are relatively small. Thus, the corresponding $\Delta\Delta G^{\circ}_{pK_a}$ values (in kJ mol^{-1}) [139,140] of ~5.1 and 1.1 for **8** and **7**, respectively, have been attributed to the variable strength of electrostatic interactions between the electron densities of the involved atoms in the offset stacked nucleobases as well as with that of the phosphates across the ssRNA strand. Studies with proteins (see the section "pK_a perturbation in proteins") have predominantly showed the perturbation [108–110] with lowering of pK_a, however, few studies have also reported elevated pK_a. Thus, our experimental evidences of apparent pK_a–2 (Type 2 effect) (i.e., elevated pK_a of 9-guaninyl from the marker protons of neighboring nucleobases in a sequence context) have shown new aspects of local electrostatic modulation via nearest-neighbor interactions in single-stranded nucleic acids.

Propagation of electrostatic interplay across the single-stranded RNA chain

It has been observed [138] that the propagation of the interplay of various electrostatic interaction among offset stacked nucleobases across the hexameric oligo-ssRNA chain indeed extends (NMR detectable) from the first to the sixth nucleotide with maximal effect until the third nucleobase in the hexameric ssRNA (**8**). This demonstrates that each nucleobase across the oligo-chain is engaged through a variable electrostatic interaction with the next neighbor(s), depending upon their individual pseudoaromatic characters modulated by their respective microenvironments.

It has also been demonstrated [138] that the pseudoaromatic nucleobases in the hexameric RNA **8** as well as other oligo-ssRNAs **5**–**7** constitute an electronically coupled heterocyclic system right across the pH range, 6.7 to 12.1. The specific generation of a single guanylate ion in the hexameric RNA molecule allowed us to demonstrate that the electrostatic atom-$\pi\sigma$ interaction indeed extends from the first to the sixth nucleotide involving each nucleotide step in a stepwise manner (ca. 3.4 Å) in a single-stranded hexameric RNA in the neutral state. This also shows that the strength of the stabilizing stacking interaction is strongest under the quasi-physiological condition at the neutral state.

The stability of the stacked helical ssRNA conformation is reduced by the destabilizing anion(G^-)-π/dipole($Im^{\delta-}$) interaction as a result of the generation of the 9-guanylate ion as observed [138] from relative chemical shift, oligomerization shift as well as sugar conformation analyses in **5**–**8**. This destabilizing effect in the deprotonated RNA becomes less pronounced as the ssRNA chain length increases because of opposing atom-$\pi\sigma$ interaction (major) as well as minor anion(G^-)-π/dipole($Py^{\delta+}$) interactions.

The magnitude of the pH-dependent chemical shift change [see Fig. 7 for $\Delta\delta_{(N-D)}$ for tri-, tetra-, penta-, and hexameric ssRNAs (**5a**, **5b**, **6**–**8**)] in any of the aromatic protons in either of the two coupled nucleobases differs in a variable manner [138] depending upon the geometry of stacking, electron density around the heteroatom as well as the sequence context (Fig. 8). Thus, the physicochemical character (i.e., the cross-modulation of pseudoaromaticity) of an individual nucleobase in an oligonucleotide is determined in a tunable manner, depending both upon the geometry and the strength of the nearest-neighbor interaction.

The importance [138] of generating single negatively charged center, upon N^1 deprotonation of 9-guaninyl residue in oligo-ssRNAs, is that the nucleobases across the strand become partially electronically decoupled (destacked) because of Coulombic repulsion, which is evidenced by the destablization of the helix, compared to that in the neutral state. As the pH becomes alkaline, the phosphate being negatively charged repels the negatively charged guanylate anion. The alignment of all other nucleobases in the ssRNA sequence is then dictated by the new stacking orientation and/or planar nucleobase rotation due to anion(G^-)-π/dipole($Im^{\delta-}$) interaction and guanylate-phosphate repulsion. This is then modulated in a variable manner depending upon the pseudoaromatic character of the next neighboring nucleobase and propagating further across the RNA chain in a stepwise manner, thereby steering the two-state stacking $\leftrightarrows$ destacking equilibria to a relatively more destacked state, causing a destabilization of the ssRNA helix.

CONCLUSION

1. Thus, obtaining the pK_as of a single ionization point from all marker protons of each nucleotide residue in the vicinity allows us to experimentally examine the microscopic change of the electronic environment around each constituent nucleobase along the RNA chain in a stepwise manner. The net result of this electrostatic cross-talk between two electronically coupled neighboring aglycones as a result of base-base stacking is creation of a *unique set of aglycones* in an oligo or polynucleotide, whose physicochemical properties are completely dependent upon the nearest-neighbor electrostatic interactions. This has considerable implication in the specific ligand bind-

ing ability like in aptamer recognition, RNA catalysis, and most probably in codon-anticodon interaction.

2. The pH titration study offers an in depth understanding of the nature of the electrostatic mediated self-assembly process by simple intramolecular stacking interactions and the conformational dynamics in the single-stranded RNA, which are normally very difficult to measure quantitatively by state-of-the-art NMR spectroscopy. We have thus demonstrated [60,136–138] above that a simple 1D NMR-based pH-dependent titration profile can be used as a major experimental tool to identify the nature of nearest-neighbor aromatic interaction provided the complex has a specific *single* protonation or deprotonation site over the pH range studied, as designed in our ssRNA sequences. Interestingly, such a titration method also gives an enormous insight into the nature of weak noncovalent contribution in such aromatic interactions by providing the energetics (from pK_a perturbation owing to nearest-neighbor interactions), geometry (from relative chemical shift change as a function of pH) of such interactions, which are of great importance in molecular recognition pattern as well as the self-assembly process in both biological as well as nonbiological system.
3. The generation of a new anionic center in the oligo-ssRNAs destabilizes the stacked state in a distance-dependent manner, which can be thermodynamically described using our pH titration procedure. Thus, pH-titration study with NMR in conjunction with structure elucidation by NMR/ab initio or X-ray and subsequent Poisson–Boltzmann calculation of the surface potential distribution may allow us to map the overall electrostatic effect in ssRNA, in general. This may help us to understand why the sequence context is so important for biological recognition, interaction, and function of RNA in general.
4. The sequence-dependent modulation of the pseudoaromatic character of the nucleobases in an oligo-RNA would change the ligand binding properties both by weak interactions (electrostatic, hydrophobic, van der Waals) as well as by hydrogen-bonding interactions as found in the aptamers. However, we envision that the spread of these electrostatic interactions along the RNA chain would depend upon whether the neighboring nucleobases are electronically coupled owing to offset stacking or not (ON-OFF switch owing to the stacking $\leftrightarrows$ destacking equilibrium).
5. Thus, in an RNA sequence, $P^1Q^1\underline{\mathbf{N}}Q^2P^2$, the actual physicochemical integrity of $\underline{\mathbf{N}}$ 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, $\underline{\mathbf{N}}$. Thus, the pseudoaromatic character of $\underline{\mathbf{N}}$ can have at least 2^4 numbers of variations, depending upon the chemical nature of the neighboring Q^1 and Q^2, which therefore implies that a given nucleobase sequence in a polynucleotide chain constitutes a unique extended genetic code, which can be turned ON or OFF depending upon the intrinsic dynamics of folding and unfolding within the molecule owing to the sequence context or interaction with an external ligand.

ACKNOWLEDGMENTS

Generous financial support from the Swedish Natural Science Research Council (Vetenskapsrådet), the Stiftelsen för Strategisk Forskning, and Philip Morris, Inc., USA is gratefully acknowledged.

REFERENCES

1. W. Saenger. *Principles of Nucleic Acid Structure*, Springer Verlag, Berlin (1988).
2. V. A. Bloomfield, D. M. Crothers, I. Tinoco. *Nucleic Acids: Structures, Properties and Functions*, University Science Books, Sausalito, CA (1999).

3. C. Thibaudeau and J. Chattopadhyaya. *Stereoelectronic Effects in Nucleosides and Nucleotides and their Structural Implications*, Department of Bioorganic Chemistry, Uppsala University Press (jyoti@boc.uu.se), Sweden (1999).
4. J. Plavec, C. Thibaudeau, J. Chattopadhyaya. *Pure Appl. Chem.* **68**, 2137 (1996).
5. P. Acharya, J. Issakson, P. I. Pradeepkumar, J. Chattopadhyaya. *Collect. Czech. Chem. Commun., Collection Symposium Series* (Chemistry of Nucleic Acid Components) **5**, 99 (2002).
6. S. I. Chan and J. H. Nelson. *J. Am. Chem. Soc*. **91**, 168 (1969).
7. C. Altona. In *Structure and Conformation of Nucleic Acids and Protein- Nucleic Acid Interactios*, M. Sundaralingam and S. T. Rao (Eds.), p. 613, University Park Press, Baltimore, MD, USA (1975).
8. C.-H. Lee, F. S. Ezra, N. S. Kondo, R. H. Sarma, S. Danyluk. *Biochemistry* **15**, 3627 (1976).
9. C. S. M. Olsthoorn, L. J. Bostelaar, J. F. M. de Rooij, J. H. van Boom. *Eur. J. Biochem.* **115**, 309 (1981).
10. H. Simpkins and E. G. Richards. *Biochemistry* **6**, 2513 (1967).
11. R. Luo, H. S. R. Gilson, J. Potter, M. K. Gilson. *Biophys. J.* **80**, 140 (2001).
12. M. M. Warshaw and I. Tinoco, Jr. *J. Mol. Biol.* **13**, 54 (1965).
13. R. A. Friedman and B. Honig. *Biophys. J.* **69**, 1528 (1995).
14. N. S. Kondo and S. S. Danyluk. *Biochemistry* **15**, 756 (1976).
15. A. D. Broom, M. P. Schweizer, P. O. P. Ts'o. *J. Am. Chem. Soc*. **89**, 3612 (1967).
16. N. H. Kolodny and A. C. Neville. *Biopolymers* **19**, 2223 (1980).
17. M. E. Burkard, R. Kierzek, D. H. Turner. *J. Mol. Biol.* **290**, 967 (1999).
18. J. Kim, A. E. Walter, D. H. Turner. *Biochemistry* **35**, 13753 (1996).
19. S. Bommarito, N. Peyret, J. SantaLucia, Jr. *Nucleic Acids Res.* **28**, 1929 (2000).
20. H. Rosemeyer and F. Seela. *J. Chem. Soc., Perkin Trans. 2* 746 (2002).
21. T. Ohmichi, S-i. Nakano, D. Miyoshi, N. Sugimoto. *J. Am. Chem. Soc*. **124**, 10367 (2002).
22. C. A. Hunter and J. K. M. Sanders. *J. Am. Chem. Soc*. **112**, 5525 (1990).
23. C. A. Hunter, K. R. Lawson, J. Perkins, C. J. Urch. *J. Chem. Soc., Perkin Trans.* **2**, 651 (2001).
24. W. B. Jennings, B. M. Farrell, J. F. Malone. *Acc. Chem. Res.* **34**, 885 (2001).
25. M. Waters. *Curr. Opin. Chem. Biol.* **6**, 736 (2002).
26. F. Cozzi, R. Annuziata, M. Benaglia, M. Cinquini, L. Rainmondi, K. K. Baldridge, J. S. Siegel. *Org. Biomol. Chem.* **1**, 157 (2003).
27. E. A. Meyer, R. K. F. Castellano, Diederich. *Angew. Chem., Int. Ed.* **42**, 1210 (2003).
28. E. T. Kool. *Annu. Rev. Biophys. Biomol. Struct.* **30**, 1 (2001).
29. J. West, S. Mecozzi, D. A. Dougherty. *J. Phys. Org. Chem.* **10**, 347 (1997).
30. E. Kim, S. Paliwal, C. S. Wilcox. *J. Am. Chem. Soc*. **120**, 11192 (1998).
31. S. Paliwal, S. Geib, C. S. Wilcox. *J. Am. Chem. Soc*. **116**, 4497 (1994).
32. F. Cozzi, M. Cinquini, R. Annuziata, J. S. Siegel. *J. Am. Chem. Soc*. **114**, 5330 (1993).
33. F. Cozzi, M. Cinquini, D. T. R. Annuziata, J. S. Siegel. *J. Am. Chem. Soc*. **114**, 5729 (1992).
34. Y. Umezawa, S. Tsuboyama, H. Takahashi, J. Uzawa, M. Nishio. *Tetrahedron* **55**, 10047 (1999).
35. H. Suezawa, T. Hashimoto, K. Tsuchinaga, T. Yoshida, T. Yuzuri, K. Sakakibara, M. Hirota, M. J. Nishio. *J. Chem. Soc., Perkin Trans. 2*, 1243 (2000).
36. D. A. Doughterty and D. A. Stauffer. *Science* **250**, 1558 (1990).
37. J. C. Ma and D. A. Doughterty. *Chem. Rev.* **97**, 1303 (1997).
38. C. Garau, D. Quinonero, A. Frontera, P. Ballester, A. Costa, P. M. Deya. *New J. Chem.* **27**, 211 (2003).
39. D. Quinonero, C. Garau, C. Rotger, A. Frontera, P. Ballester, A. Costa, P. M. Deya. *Angew. Chem., Int. Ed.* **41**, 3389 (2002).
40. P. Gale, K. Navakhun, S. Camiolo, M. E. Light, M. Hursthouse. *J. Am. Chem. Soc*. **124**, 11228 (2002).
41. M. Mascal, A. Armstrong, M. D. Bartberger. *J. Am. Chem. Soc*. **124**, 6274 (2002).

42. C. A. Hunter. *J. Mol. Biol.* **230**, 1025 (1993).
43. J. Ribas, E. Cubero, J. Luque, M. Orozco. *J. Org. Chem.* **67**, 7057 (2002).
44. S. B. Ferguson, E. M. Seward, F. Diederich, E. M. Sanford, A. Chou, P. Inocencio-Szweda, C. B. Knobler. *J. Org. Chem.* **53**, 5595 (1988).
45. L. Nicolas, M. Beugelmans-Verrier, J. Guilhem. *Tetrahedron* **37**, 3847 (1981).
46. D. A. Stauffer, R. E. Barrans, Jr., D. A. Doughterty. *Angew. Chem., Int. Ed.* **29**, 915 (1990).
47. T. Ishida, M. Shibata, K. Fuji, M. Inoue. *Biochemistry* **22**, 3571 (1983).
48. J. Landauer and H. McConnell. *J. Am. Chem. Soc*. **74**, 1221 (1952).
49. A. M. Z. Slawin, N. Spencer, J. F. Stoddart, D. J. Williams. *J. Chem. Soc., Chem. Commun.* 1070 (1987).
50. L. F. Newcomb and S. H. Gellman. *J. Am. Chem. Soc*. **116**, 4993 (1994).
51. M. J. Rashkin and M. L. Waters. *J. Am. Chem. Soc*. **124**, 1860 (2002).
52. G. A. Breault, C. A. Hunter, P. C. Mayers. *J. Am. Chem. Soc*. **120**, 3402 (1998).
53. T. S. Snowden, A. P. Bisson, E. V. Anslyn. *J. Am. Chem. Soc*. **121**, 6324 (1999).
54. A. J. Goodman, E. C. Breinlinger, C. M. McIntosh, L. N. Grimaldi, V. M. Rotello. *Org. Lett.* **3**, 1531 (2001).
55. A. S. Shetty, J. S. Zhang, J. S. Moore. *J. Am. Chem. Soc*. **118**, 1019 (1996).
56. G. J. Gabriel and B. L. Iversson. *J. Am. Chem. Soc*. **124**, 15174 (2002).
57. F. J. Carver, C. A. Hunter, E. M. Seward. *J. Chem. Soc., Chem. Commun.* **775**, (1998).
58. H. Adams, F. J. Carver, C. A. Hunter, J. C. Morales, E. M. Seward. *Angew. Chem., Int. Ed.* **35**, 1542 (1996).
59. S. Yamada and C. Morita. *J. Am. Chem. Soc*. **124**, 8184 (2002).
60. P. Acharya, O. Plashkevych, C. Morita, S. Yamada, J. Chattopadhyaya. *J. Org. Chem.* **68**, 1529 (2003).
61. M. J. Packer, M. P. Dauncey, C. A. Hunter. *J. Mol. Biol*. **295**, 71 (2000).
62. M. J. Packer and C. A. Hunter. *J. Am. Chem. Soc*. **123**, 7399 (2001).
63. P. Hsu, M. R. Hodel, W. J. Thomas, L. J. Talyor, C. H. Hagedorn, A. E. Hodel. *Biochemistry* **39**, 13730 (2000).
64. G. Hu, P. D. Gershon, A. E. Hodel, F. A. Quiocho. *Proc. Natl. Acad. Sci. USA* **96**, 7149 (1999).
65. G. Hu, A. Oguro, C. Li, P. D. Gershon, F. A. Quiocho. *Biochemistry* **41**, 7677 (2002).
66. I. Berger, M. Egli, A. Rich. *Proc. Natl. Acad. Sci. USA* **93**, 12116 (1996).
67. S. Sarkhel, A. Rich, M. Egli. *J. Am. Chem. Soc*. **125**, 8998 (2003).
68. M. Rooman, J. Liévin, E. Buisine, R. Wintjens. *J. Mol. Biol.* **319**, 67 (2002).
69. R. Wintjens, J. Liévin, M. Rooman, E. Buisine. *J. Mol. Biol.* **302**, 395 (2000).
70. L. McFail-Isom, X. Shui, L. D. Williams. *Biochemistry* **37**, 17105 (1998).
71. Y. Umezawa and M. Nishio. *Nucleic Acids Res*. **30**, 2183 (2002).
72. N. Zacharias and D. A. Dougherty. *Trends Pharmacol. Sci.* **23**, 281 (2002).
73. D. D. Boehr, A. R. Farley, G. D. Wright, J. R. Cox. *Chem. Biol*. **9**, 1209 (2002).
74. C. D. Tatko and M. L. Waters. *J. Am. Chem. Soc*. **124**, 9372 (2002).
75. S. M. Butterfield, P. R. Patel, M. L. Water. *J. Am. Chem. Soc*. **124**, 9751 (2002).
76. L. K. Tsou, C. D. Tatko, M. L. Waters. *J. Am. Chem. Soc*. **124**, 14917 (2002).
77. F. L. Gervasio, R. Chelli, P. Procacci, V. Schettino. *Proteins: Struct., Funct., Genet.* **48**, 117 (2002).
78. Z. Zhou and R. P. Swenson. *Biochemistry* **35**, 15980 (1996).
79. J. P. Gallivan and D. A. Dougherty. *Proc. Natl. Acad. Sci. USA* **96**, 9459 (1999).
80. C. Biot, E. Buisine, J.-M. Kwasigroch, R. Wintjens, M. Rooman. *J. Biol. Chem.* **277**, 40816 (2002).
81. A. Fersht. *Enzyme Structure and Mechanism*, W. H. Freeman, New York (1984).
82. G. J. Connell and M. Y. Yarus. *Science* **264**, 1137 (1994).
83. P. Legault and A. Pardi. *J. Am. Chem. Soc*. **116**, 8390 (1994).

84. P. Legault and A. Pardi. *J. Am. Chem. Soc*. **119**, 6621 (1997).
85. S. Ravindranathan, S. E. Butcher, J. Feigon. *Biochemistry* **39**, 16026 (2000).
86. Z. Cai and I. Tinoco, Jr. *Biochemistry* **35**, 6026 (1996).
87. G. W. Muth, L. Ortoleva-Donnelly, S. A. Strobel. *Science* **289**, 947 (2000).
88. L. Xiong, N. Polacek, P. Sander, E. C. Böttger, A. Mankin. *RNA* **7**, 1365 (2001).
89. A. C. Drohat and J. T. Stivers. *J. Am. Chem. Soc*. **122**, 1840 (2000).
90. A. C. Drohat and J. T. Stivers. *Biochemistry* **39**, 11865 (2000).
91. S. Nakano, D. M. Chadelavada, P. C. Bevilacqua. *Science* **287**, 1493 (2000).
92. A. Lupták, A. R. Ferré-D'Amaré, K. Zhou, K. W. Zilm, J. A. Doudna. *J. Am. Chem. Soc*. **123**, 8447 (2001).
93. A. T. Perrotta, I. Shih, M. D. Been. *Science* **286**, 123 (1999).
94. Y. Boulard, J. A. H. Cognet, J. Gabarro-Arpa, M. LeBret, L. C. Sowers, G. V. Fazakerley. *Nucleic Acids Res.* **20**, 1933 (1992).
95. C. Wang, H. Gao, B. L. Gaffney, R. A. Jones. *J. Am. Chem. Soc*. **113**, 5486 (1991).
96. P. C. Bevilacqua. *Biochemistry* **42**, 2259 (2003).
97. G. J. Narlikar and D. Herschlag. *Annu. Rev. Biochem.* **66**, 19 (1997).
98. S.-O. Shan and D. Herschlag. *Proc. Natl. Acad. Sci. USA* **93**, 14474 (1996).
99. A. Yoshida, S. Shuou, D. Herschlag, J. A. Piccirilli. *Chem. Biol.* **7**, 85 (2000).
100. D. Herschlag, F. Eckstein, T. R. Cech. *Biochemistry* **32**, 8312 (1993).
101. J. Sancho, L. Serrano, A. R. Fersht. *Biochemistry* **31**, 2253 (1992).
102. M. Tollinger, K. A. Crowhurst, L. E. Kay, J. D. Forman-Kay. *Proc. Natl. Acad. Sci. USA* **100**, 4545 (2003).
103. H. A. Chen, M. Pfuhl, M. S. B. McAlister, P. C. Driscoll. *Biochemistry* **39**, 6814 (2000).
104. K. Langsetmo, J. A. Fuchs, C. Woodward. *Biochemistry* **30**, 7603 (1991).
105. B. Ibarra-Molero, V. V. Loladze, G. I. Makhatadze, J. M. Sanchez-Ruiz. *Biochemistry* **38**, 8138 (1999).
106. N.-C. Ha, M.-S. Kim, W. Lee, K. Y. Choi, B.-H. Oh. *J. Biol. Chem.* **275**, 41100 (2000).
107. W. Schaller and A. D. Robertson. *Biochemistry* **34**, 4714 (1995).
108. J. Song, M. Laskowski, M. A. Qasim, J. L. Markley. *Biochemistry* **42**, 2847 (2003).
109. R. Consonni, I. Arosio, B. Belloni, F. Fogolari, P. Fusi, E. Shehi, L. Zetta. *Biochemistry* **42**, 1421 (2003).
110. M. D. Joshi, G. Sidhu, J. E. Neilsen, G. D. Brayer, S. G. Withers, L. P. McIntosh. *Biochemistry* **40**, 10115 (2001).
111. E. J. Petersson, A. Choi, D. S. Dahan, H. A. Lester, D. A. Dougherty. *J. Am. Chem. Soc*. **124**, 12662 (2002).
112. S. Limmer, H. P. Hofmann, G. Ott, M. Sprinzl, *Proc. Natl. Acad. Sci. USA* **90**, 6199 (1993).
113. D. Ayer and M. Yarus. *Science* **231**, 393 (1986).
114. Z. Du, D. P. Giedroc, D. W. Hoffmann. *Biochemistry* **35**, 4187 (1996).
115. J. Zhu and R. M. Wartell. *Biochemistry* **36**, 15326 (1997).
116. J. Zhu and R. M. Wartell. *Biochemistry* **38**, 15986 (1999).
117. E. M. Anderson, W. A. Halsey, D. S. Wuttke. *Biochemistry* **42**, 3751 (2003).
118. T. Nishinaka, Y. Ito, S. Yokoyama, T. Shibata. *Proc. Natl. Acad. Sci. USA* **94**, 6623 (1997).
119. J.-J. Toulmé. NATO ASI Series, Series A, Life Sci. 101 (Chromosomal Protein Gene Expression), 263 (1985).
120. G. Vesnaver and K. J. Breslauer. *Proc. Natl. Acad. Sci. USA* **88**, 3569 (1991).
121. M. Weinfeld, K.-J. M. Soderlind, G. W. Buchko. *Nucleic Acids Res.* **21**, 621 (1993).
122. R. Bar-Ziv and A. Libchaber. *Proc. Natl. Acad. Sci. USA* **98**, 9068 (2001).
123. O. Igoucheva, V. Alexeev, K. Yoon. *Gene Therapy* **8**, 391 (2001).
124. R. I. M. Wadsworth and M. F. White. *Nucleic Acids Res.* **29**, 914 (2001).
125. J. Ren and J. B. Chaires. *Biochemistry* **38**, 16067 (1999).

126. E. Bochkareva, V. Belegu, S. Korolev, A. Bochkarev. *EMBO J.* **20**, 612 (2001).
127. D. P. Aalberts, J. M. Parman, N. L. Goddard. *Biophys. J.* **84**, 3212 (2003).
128. X. Li and D. R. Liu. *J. Am. Chem. Soc*. **125**, 10188 (2003).
129. S. Limmer, H. P. Hofmann, G. Ott, M. Sprinzl. *Proc. Natl. Acad. Sci. USA* **90**, 6199 (1993).
130. S. K. Swamynathan, A. Nambiar, R. V. Guntaka. *FASEB J.* **12**, 515 (1998).
131. A. Antson. *Curr. Opin. Struct. Biol.* **10**, 87 (2000).
132. N. Handa, O. Nureki, K. Kurimoto, I. Kim, H. Sakamoto, Y. Shimura, Y. Muto, S. Yokoyama. *Nature* **398**, 579 (1999).
133. R. C. Deo, J. B. Bonanno, N. Sonenberg, S. K. Burley. *Cell* **98**, 835 (1999).
134. C. E. Bogden, D. Fass, N. Bergman, M. D. Nichols, J. M. Berger. *Mol. Cell* **3**, 487 (1999).
135. A. Antson, E. J. Dodson, G. Dodson, R. B. Greaves, X. P. Chen, P. Gollnick. *Nature* **401**, 235 (1999).
136. S. Acharya, P. Acharya, A. Földesi, J. Chattopadhyaya. *J. Am. Chem. Soc*. **124**, 13722 (2002).
137. P. Acharya, S. Acharya, A. Földesi, J. Chattopadhyaya. *J. Am. Chem. Soc*. **125**, 2094 (2003).
138. P. Acharya, S. Acharya, N. V. Amirkhanov, P. Cheruku, A. Földesi, J. Chattopadhyaya. *J. Am. Chem. Soc.* **125**, 9948 (2003).
139. D. D. Perrin, B. Dempsey, E. P. Serjeant. *pK_a Prediction for Organic Acids and Bases*, Chapman and Hall, New York (1981).
140. K. A. Sharp and B. Honig. *Annu. Rev. Biophys. Chem.* **19**, 301 (1990).