ARTICLE

OBC
www.rsc.org/obc

The 5-Me of thyminyl (T) interaction with the neighboring nucleobases dictate the relative stability of isosequential DNA–RNA hybrid duplexes†

Subhrangsu Chatterjee, Wimal Pathmasiri and Jyoti Chattopadhyaya*

Department of Bioorganic Chemistry, Box 581, Biomedical Center, Uppsala University, SE-75123, Uppsala, Sweden. E-mail: jyoti@boc.uu.se; Fax: +46 18554495

Received 5th August 2005, Accepted 5th September 2005
First published as an Advance Article on the web 29th September 2005

Relative energetic contributions from the base-pairing [ΔG°_{bp}] vis-a-vis stacking [$\Delta G^{\circ}_{stacking}$] to the total free-energy of stabilization [ΔG°_{37}] for 14 pairs of isosequential hybrid DNA–RNA duplexes (taken from E. A. Lesnik and S. M. Freier, *Biochemistry* ,1995, 34, 10807) have been dissected in order to understand the differences in the intrinsic nature of the electrostatic forces that are responsible for the self-assembly of the heteroduplexes compared to homoduplexes. The pK_a differences between the monomeric nucleotide 3′-ethylphosphates [(d/rN)pEt] as well as nucleotide 3′,5′-bis-ethylphosphates [Etp(d/rN)pEt] in both 2′-deoxy (dN) and ribo (rN) series (N = A/G/C/T/U), as the model donor and acceptor (in which stacking is completely eliminated) mimicking those of the internucleotide monomer building blocks of a duplex, can be qualitatively used (P. Acharya, P. Cheruku, S. Chatterjee, S. Acharya and J. Chattopadhyaya, *J. Am. Chem. Soc.*, 2004, 126, 2862) to understand the strength of base-pairing energies in different DNA–RNA (DR), RNA–DNA (RD), DNA–DNA (DD), and RNA–RNA (RR) duplexes. The study has led us to show the following. (1) As the number of excess %T in DR duplexes compared to the isosequential RD duplexes increase the differences in their thermal stabilization [ΔT_m]$_{DR-RD}$ increase and *vice-versa* (2) The total relative stabilizations, [$\Delta\Delta G^{\circ}_{37}$]$_{DR-RD}$ among the 14 pairs of isosequential DR and RD duplexes (E. A. Lesnik and S. M. Freier, *Biochemistry* ,1995, 34, 10807) are wholly dependent on the differences in the number of 5-Me(T) stacking interactions with the nearest-neighbors in the D strands of DR duplexes compared to that of the RD duplexes (3) In the relative stabilization of the DR or RD duplexes differences in the free-energy of stackings [$\Delta\Delta G^{\circ}_{stacking}$]$_{DR-RD}$ play a more significant role than the differences in the free-energy of base-pairing, [$\Delta\Delta G^{\circ}_{bp}$]$_{DR-RD}$. In contradistinction, our experimental data shows that RNA–RNA duplexes are more stable than DNA–DNA duplexes because of larger energy gain from the base-pairing in the former compared to the latter (P. Acharya, P. Cheruku, S. Chatterjee, S. Acharya and J. Chattopadhyaya, *J. Am. Chem. Soc.*, 2004, 126, 2862).

Introduction

The RNA–RNA (**RR**) duplexes are more stable than DNA–DNA (**DD**)/DNA–RNA (**DR**)/RNA–DNA (**RD**) duplexes[1] with the identical sequence context. The actual order of the thermodynamic stability of **DD**, **DR** and **RD** however depends upon their sequence composition.[2–6] Understanding of factors governing the stability of the heteroduplexes (**DR** *versus* **RD**) are important because they occur in cells as a result of transcription,[7] in Okazaki fragments in replication,[8] and in reverse transcription when cells are infected by retroviruses[9] or in the down-regulation of genes by inhibition of mRNAs[10–13] using the antisense strategy. In homoduplexes, **DD** and **RR**, two types of base-pairings are possible, *i.e.* rA–U/rG–rC in **RR** and dA–dT/dG–dC in **DD**. In **DR** and **RD** heteroduplexes, however, four types of base-pairings are possible depending upon the location of the nucleobases in the DNA or in the RNA strand, [*i.e.* (1) dG–rC, (2) dC–rG, (3) dT–rA, and (4) dA–rU in **DR**, and (1) rG–dC, (2) rC–dG, (3) rA–dT, (4) rU–dA in **RD**].

On the basis of the pK_a differences between the model donor and acceptor of the monomeric nucleotide 3′-ethylphosphates [(d/rN)pEt] as well as nucleotide 3′,5′-bis-ethylphosphates [Etp(d/rN)pEt] in both 2′-deoxy (dN) and ribo (rN) series, which are completely devoid of internucleotidic base–base stacking,[14] we have demonstrated that the hydrogen bondings in rA–rU and rG–rC base-pairs in **RR** duplexes are more stable than that of the dA–dT and dG–dC base-pairs[14] in **DD** duplexes by 4.3 and 1 kJ mol^{-1}, respectively, which have been also confirmed by TROSY-HSQC experiments,[15] as well as by imino-proton exchange rate of G–C base-pair flanking the (A–U/T) pair.[16] Additional stabilization in **RR** duplexes by 2′OH(n)–O4′(n + 1) hydrogen bonding has also been shown to play a key role in the X-ray structure of hammerhead ribozyme[17] as well as in the MD simulation study.[18] Kool and coworkers have shown that the effects of thyminyl-methyl and the 2′-OH group of ribose are independent of one another and that the C5-methyl group in pyrimidines are in all cases stabilizing, while 2′-OH groups can be stabilizing or destabilizing, depending on the type of double or triple helical nucleic acid complexes.[19]

Crystal structures of various TATA box binding proteins (TBP) bound to their promoter DNA (TATA box: $[^{3'-TATA-5'}_{5'-ATAT-3'}]$[20]) showed a number of CH–C_{sp2}(π) contacts between the nucleotides adjacent to each other within the same strand at the boundary of TBP and the TATA box minor groove,[20] which consequently forms very stable robust AT tract.[20–22] Recent crystal structure analysis of TATA boxes, $[^{3'-TATA-5'}_{5'-ATAT-3'}]$,[21] within DD duplexes have also clearly shown that the C3 symmetric 5-methyl group in the 1-thyminyl moiety favorably interacts with the π cloud of the fused imidazole ring of the 9-adeninyl base at the 5′-end [Me(T)–π(A)interaction]. This is consistent with the fact that C5 methyl group of 1-(5-methylcytosinyl)[23] or C5-propynyl derivative[23] of 1-uracilyl in the D strand stabilizes the **DR** duplex by ~2.5 °C per modification compared to native 1-cytosinyl. These interactions[20,21] between the C5-methyl group of 1-thyminyl and the π cloud of the preceding nucleobase at the 5′-end indeed is also known to stabilize the intrastrand stacking in the heteroduplexes ("twin A/T–Me interaction"),[21,24,25] the basis for the construction of DNA curvature[26] and bending.[27]

† Electronic supplementary information (ESI) available: Additional tables of calculations. See DOI: 10.1039/b511139k

DOI: 10.1039/b511139k

In our earlier study[14] we have shown that RNA–RNA duplex is more stabilized than the isosequential DNA–DNA (DD) duplex because of stronger hydrogen bonding in the rC–rG and rA–rU base-pairing in comparison with the dC–dG and dA–dT base-pairings than that of the stacking interaction. We here show, in contradistinction, that in case of isosequential DR and RD duplexes the main contribution to ΔG°_{37} arises from the contribution of the stacking energies, not from the base-pairing energies. We here also show that it is the 5-methyl group of T interactions with the neighboring bases that is the most important contributor than that of the other stacking interactions in the overall stacking-promoted stabilization of the heteroduplexes. We have come to this conclusion by comparing the effect of excess number of T in the D strand by pairwise (14 pairs) subtraction of T residues among the isosequential DR and RD duplexes.

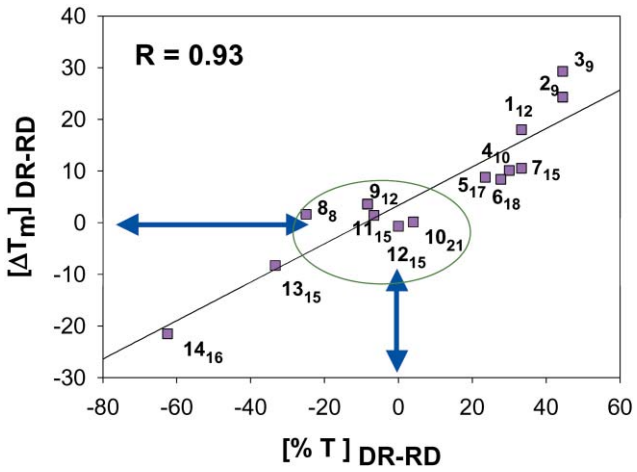


Fig. 1 A plot of the relative thermal stability, $[\Delta T_m]_{DR\text{–}RD}$, of isosequential **DR** compared to **RD** duplexes as a function of increase of excess %T in **DR** with respect to the **RD**, $[\%T]_{DR\text{–}RD}$. The chain length of all the sequences, **1**–**14** (Table 1), are marked as the subscript in the correlation plots (1–3). Note when the number of T is the same in **DR** and **RD** duplexes (*i.e.* $[\%T]_{DR\text{–}RD} = 0$, shown by arrow and ellipse) their corresponding thermal stabilities, $[\Delta T_m]_{DR\text{–}RD}$, are also very close to zero.

Results and discussion

From the T_m analysis of the isosequential **DR** and **RD** duplexes (**1–14**)[4] (Table 1 and its legend), it can be seen that the relative stability of the **DR** duplex compared to the isosequential **RD** duplexes, $[\Delta T_m]_{DR\text{–}RD}$, increases as the excess %T residues in **DR** increases compared to **RD**, $[\%T]_{DR\text{–}RD}$, and *vice versa*. Thus, a linear regression plot of $[\Delta T_m]_{DR\text{–}RD}$ *versus* $[\%T]_{DR\text{–}RD}$ shows (Fig. 1) that as the excess %T moieties in the **DR** heteroduplex increases over the **RD**, the relative stability of the former increases in comparison with the latter [Pearson's correlation coefficient, $R = 0.93$]. This can be more clearly seen upon inspection of the plot in Fig. 1, when the number of T's is the same in DR and RD duplexes (*i.e.* $[\%T]_{DR\text{–}RD} = 0$, shown by arrow) their corresponding thermal stabilities, $[\Delta T_m]_{DR\text{–}RD}$, are also very close to zero. This means that when $[\%T]_{DR\text{–}RD}$ is larger than zero, the $[\Delta T_m]_{DR\text{–}RD}$ is also larger than zero, and *vice versa*.

Table 1 The relative T_m stabilities of heteroduplexes (**DR**/**RD**)[4], $\Delta T_{m[DR\text{–}RD]}$, is related to their differences in %T-content, $[\%T]_{DR\text{–}RD}$. and quantification (kcal mol^{-1}) of the energetic contributions from the relative base-pairing $[\Delta\Delta G^{\circ}_{bp}]_{DR\text{–}RD}$, and stacking $[\Delta\Delta G^{\circ}_{stacking}]_{DR\text{–}RD}$ in $[\Delta G^{\circ}_{37}]_{DR\text{–}RD}$

			Eqn (1)			Eqn (2c)	Eqn (3)	
			$[\Delta\Delta G^{\circ}_{37}]_{DR\text{–}RD}$[c]/kcal mol^{-1}					
Sequence[a]	Duplex sequences[a]	$\Delta T_{m\,[DR\text{–}RD]}$[b]	DR	RD	DR–RD	$[\Delta\Delta G^{\circ}_{bp}]_{DR\text{–}RD}$[d]	$[\Delta\Delta G^{\circ}_{stacking}]_{DR\text{–}RD}$[e]	$[\%\,T]_{DR\text{–}RD}$[f]
(1_{DR}–1_{RD})	5′-TCCCTCCTCTCC 3′-AGGGAGGAGAGG	18.0	−15.6	−9.6	−6.0	+0.55	−6.55	33.3
(2_{DR}–2_{RD})	5′-CCTTCCCTT 3′-GGAAGGGAA	24.3	−10.0	−5.8	−4.2	+2.16	−6.36	44.4
(3_{DR}–3_{RD})	5′-TTCCCTTCC 3′-AAGGGAAGG	29.3	−10.1	−5.1	−5.0	+2.26	−7.26	44.4
(4_{DR}–4_{RD})	5′-GCTCTCTGGC 3′-CGAGAGACCG	10.10	−11.2	−8.9	−2.3	+3.21	−5.51	30.0
(5_{DR}–5_{RD})	5′-CTCGTACCTTCCGGTCC 3′-GAGCATGGAAGGCCAGG	8.8	−17.0	−12.7	−4.3	+2.22	−6.52	23.5
(6_{DR}–6_{RD})	5′-CTCGTACCTTTCCGGTCC 3′-GAGCATGGAAAGGCCAGG	8.4	−18.1	−14.6	−3.5	+3.47	−6.97	27.7
(7_{DR}–7_{RD})	5′-TAGTTATCTCTATCT 3′-ATCAATAGAGATAGA	10.5	−10.5	−7.5	−3.0	+5.19	−8.19	33.3
(8_{DR}–8_{RD})	5′-GCACAGCC 3′-CGTGTCGG	1.6	−8.1	−7.9	−0.2	−3.67	3.47	−25.0
(9_{DR}–9_{RD})	5′-GAGCTCCCAGGC 3′-CTCGAGGGTCCG	3.6	−14.3	−13.4	−0.9	−1.87	0.93	−8.4
(10_{DR}–10_{RD})	5′GCCGAGGTCCATGTCGTACGC 3′CGGCTCCAGGTACAGCATGCG	0.1	−17.2	−18.0	+0.8	+1.26	−0.46	4.1
(11_{DR}–11_{RD})	5′-TGTACGTCACAACTA 3′-ACATGCAGTGTTGAT	1.4	−11.8	−11.2	−0.6	−2.40	1.80	−6.6
(12_{DR}–12_{RD})	5′-TATACAAGTTATCTA 3′-ATATGTTCAATAGAT	−0.7	−7.7	−7.8	+0.1	−0.57	0.67	0.0
(13_{DR}–13_{RD})	5′-CGACTATGCAAAAAC 3′-GCTGATACGTTTTTG	−8.3	−8.7	−11.3	+2.6	−7.43	10.03	−33.4
(14_{DR}–14_{RD})	5′-CGCAAAAAAAAAACGC 3′-GCGTTTTTTTTTTGCG	−21.5	−5.9	−13.0	+7.1	−13.74	20.84	−62.5

[a] All oligo sequences (**1**$_{DR}$/**1**$_{RD}$–14$_{DR}$/**14**$_{RD}$) in **DR** and **RD** combinations and their corresponding T_m data set have been taken from Lesnik. *et al.*[4] For example, in **DR**, the 5′→3′ is D and 3′→5′ is R. and for the **RD**, the 5′→3′ is R and 3′→5′ is D. All T in the R strand of **DR** or **RD** is substituted with U. [b] Thermal stability of **DR** duplexes relative to their isosequencial **RD** duplexes in °C. [c] The data for helix stability $[\Delta\Delta G^{\circ}_{37}]_{DR\text{–}RD}$ of DR (1_{DR}–14_{DR}) and RD duplexes (1_{RD}–14_{RD}) with the same sequence context used in this study are taken from ref. 4 [d] $[\Delta\Delta G^{\circ}_{bp}]_{DR\text{–}RD}$ (kcal mol^{-1}) = [{$\Sigma\,\Delta\Delta G^{\circ}_{pK_a}$ (DR) − $\Sigma\,\Delta\Delta G^{\circ}_{pK_a}$ (RD)}/4.2], [1.0 kcal mol^{-1} = 4.2 kJ mol^{-1}]. [e] $[\Delta\Delta G^{\circ}_{stacking}]_{DR\text{–}RD}$ (kcal mol^{-1}) = $[\Delta\Delta G^{\circ}_{37}]_{DR\text{–}RD} - [\Delta\Delta G^{\circ}_{bp}]_{DR\text{–}RD}$. [f] $[\%T]_{DR\text{–}RD} = [\%T]_{DR} - [\%T]_{RD}$, excess %T in **DR** in comparison with **RD**, $[\%T]_{DR}$ = [total number of T/nucleotidic chain length of the duplex]$_{DR}$ × 100. $[\%T]_{RD}$ = [total number of T/nucleotidic chain length of the duplex]$_{RD}$ × 100.

So, what is the origin of this T effect? Is it contributing toward the intrastrand stacking interactions or base-pairing? It was however clear to us from our earlier work[14] that the A:T base-pairing in DNA–DNA duplexes is weaker than the A:U base-pairing in RNA–RNA duplexes. This has led us to suspect that the stacking interaction is probably a more dominating force in the stabilization of helix in the heteroduplex, which is backed by our present analysis.

For the sake of simplicity, if we consider that there are two dominating contributors to ΔG°_{37} of a duplex (from T_m) which are H-bonding and stacking (eqn (1) and (2)). We can easily dissect these two forces by estimating the strength of H-bonding from the differences in pK_a values (ΔpK_a) between the donor and acceptor in the base-pairing.[14] Thus, we have estimated $\Delta G^{\circ}_{\text{base-pairing (bp)}}$ from all ΔpK_a[14,28,29,30] [ΔpK_a = (pK_a of the donor) − (pK_a of the acceptor)] for all A–U/T or G–C base-pairings within a duplex by summation of all $\Delta\Delta G^{\circ}_{pK_a}$ (from $\Sigma\Delta pK_a$) (eqn (2a) and (2b)). Since the strength between donor (D) and acceptor (A) in D–H···A is proportional to the pK_a difference (ΔpK_a) between the hetero atoms making the hydrogen bond,[14,28,31,32] we expected that a hydrogen bond (D–H···A ⇌ D···H–A) is to be the strongest when $\Delta pK_a = 0$,[29] *i.e.* the pK_a of the donor is the same as the pK_a of acceptor ("pK_a match").[32]

Since, in **DR** duplex, ΔpK_a for the middle (dT–rA) and (dA–rU) base-pairs are respectively 6.43 and 5.43, it signifies that dA–rU is forming a more stable hydrogen bond ($\Delta\Delta G^{\circ}_{pK_a}$ = 31.1 kJ mol^{-1}) than dT–rA ($\Delta\Delta G^{\circ}_{pK_a}$ = 36.4 kJ mol^{-1}) (see Table S1 in SI). Again, the dC–rG or rG–dC as the middle base-pair are more stable ($\Delta pK_a = 4.94$, $\Delta\Delta G^{\circ}_{pK_a} = 28.2$ kJ mol^{-1}) than those of dG–rC or rC–dG ($\Delta pK_a = 5.35$, $\Delta\Delta G^{\circ}_{pK_a} = 30.6$ kJ mol^{-1}) (see Table S1 in SI†).

Since ($\Sigma\Delta pK_a$) (Tables S3 and S4 in SI) is the sum of the total number of middle and terminal base-pairs (eqn (2b)), calculation of base-pairing energy of heteroduplexes in Table S5 in SI), of the **DR** or **RD** duplex (contributing to the total duplex stability, ΔG°_{37}), we have plotted $[\Delta G^{\circ}_{37}]_{\text{DR or RD}}$ as a function of $[\Sigma\Delta pK_a]_{\text{DR or RD}}$ (Fig. 2A and 2B and Table S3 and S4 in SI) for all 14 oligos, and performed the regression analysis and compared them with those of **DD** and **RR** duplexes (Fig. 2C and 2D and Table S3), which showed that ΔG°_{37} and the $[\Sigma\Delta pK_a]$ are well correlated for all types of homo and heteroduplexes, thereby showing that the estimation of $\Sigma\Delta pK_a$ is indeed a valid tool to dissect the energetic contribution of ΔG°_{bp} from ΔG°_{37}.

Calculation of the relative base-pairing energy of **DR** duplexes over the isosequential **RD** duplexes, $[\Delta\Delta G^{\circ}_{bp}]_{DR-RD}$, can simply be achieved by multiplying $\Sigma\Delta pK_{a[DR-RD]}$ values with $2.303RT$[14] (eqn (2a) and Table S5). Thus, the subtraction of this $[\Delta\Delta G^{\circ}_{bp}]_{DR-RD}$ from the difference of total free-energy of isosequential **DR** and **RD** duplexes, $[\Delta\Delta G^{\circ}_{37}]_{DR-RD}$, gives us directly the relative contribution of stacking, $[\Delta\Delta G^{\circ}_{Stacking}]_{DR-RD}$, (eqn (3) and Table 1).

$$[G^{\circ}_{37}] = [\Delta G^{\circ}_{\text{base-pairing (bp)}}] + [\Delta G^{\circ}_{\text{stacking}}] \quad (1)$$

$$[\Delta G^{\circ}_{\text{stacking}}] = [\Delta G^{\circ}_{37}] - [\Delta G^{\circ}_{bp}] \quad (2)$$

$$[\Delta G^{\circ}_{bp}] = 2.303RT\,pK_{a\,(\text{all donors–acceptors in the duplex})} \quad (2a)$$

$$[\Delta G^{\circ}_{bp}] = [\Sigma\Delta G^{\circ}_{pK_a\,(\text{middle bp})}] + [\Sigma\Delta G^{\circ}_{pK_a\,(\text{terminal bp})}] \quad (2b)$$

$$[\Delta\Delta G^{\circ}_{bp}]_{DR-RD} = [\Sigma\Delta\Delta G^{\circ}_{pK_a\,(\text{all bp})}]_{DR-RD} \quad (2c)$$

$$[\Delta\Delta G^{\circ}_{\text{stacking}}]_{DR-RD} = [\Delta\Delta G^{\circ}_{37}]_{DR-RD} - [\Delta\Delta G^{\circ}_{bp}]_{DR-RD} \quad (3)$$

Finally, Fig. 3A shows (eqn (2c)) that as the number of excess T ($[\%T]_{DR-RD}$) increases the relative strength of the base-pairing $[\Delta\Delta G^{\circ}_{bp}]_{DR-RD}$ (Table 1) decreases because of weaker A:T base-pairing compared to that of A:U. But in contradistinction, as the number of excess T ($[\%T]_{DR-RD}$) increases, the relative

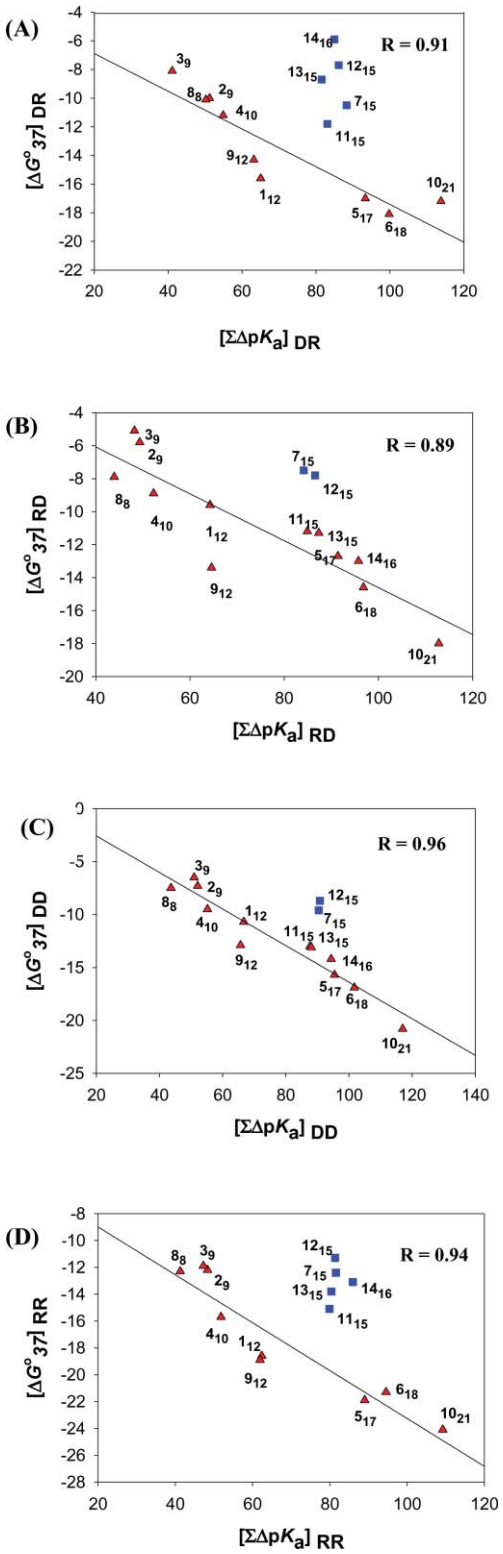
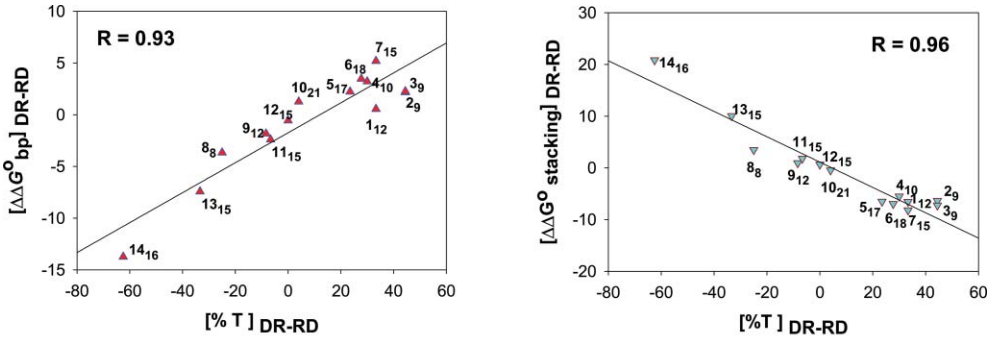


Fig. 2 Panels (A) and (B) show linear regression plots of $[\Delta G^{\circ}_{37}]_{DR}$ *versus* $(\Sigma\Delta pK_a)_{DR}$ and $[\Delta G^{\circ}_{37}]_{RD}$ *versus* $(\Sigma\Delta pK_a)_{RD}$. For details of $\Sigma\Delta pK_a$ calculation of all base-pairings, (see Table S3 in SI). Panels C and D show similar regression plots of overall free-energy of stabilization for **DD** and **RR** duplexes as a function of $\Sigma\Delta pK_a$ (*i.e.* the sum of pK_a differences between the model monomeric donors and acceptors) (Scheme 1 of ref. 14) for the sake of comparison. The TA or AT rich sequences (in blue), **7** and **12** in **RD** (Panel B) and **7**, **11**, **12**, **13**, **14** in **DR** (Panel A), **7** and **12** in **DD** (Panel C) did not show any correlation because of Me(T)–πN (N = A/G/C/T) interaction[22] which make the AT tracts robust and gives a rigid geometry. Hence these duplexes show unusual melting tendencies. Again, **7**, **11**, **12**, **13**, **14** in **RR** duplexes (Panel D) also show unusual melting tendencies because of high content of A–U base-pairings which gives relatively more rigid geometry for those **RR** duplexes compared to the other **RR**s with fewer A–U base-pairings[16] (**1–14**, marked with the chain length size as the subscript).[4]



(A) (B)

Fig. 3 Panel (A) shows that with the increase in the excess %T, $[\%T]_{DR-RD}$, in **DR** with respect to **RD** the relative base-pairing energy, $[\Delta\Delta G^{\circ}_{bp}]_{DR-RD}$, decreases (R = 0.93). Panel (B) shows that with the increase in $[\%T]_{DR-RD}$ in **DR** with respect to the **RD** the relative stacking energy, *i.e.* $[\Delta\Delta G^{\circ}_{stacking}]_{DR-RD}$ increases (R = 0.96). The $[\%\,T]_{DR-RD}$ interactions signify the sum of excess % of Me(T)–π(A/G/C/T) interactions plus the excess % of Me(T)–Me(T) interactions. All sequences with their sequence number (Table 1) and chain length as the subscript are given in the Fig. 1.

strength of the stacking forces ($[\Delta\Delta G^{\circ}_{stacking}]_{DR-RD}$) increases (eqn (3)) (Fig. 3B) because of more number of Me–T interactions with the neighboring nucleobases.

Conclusions

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

(1) One of our major conclusions is that the intrastrand stacking interaction between the C5-methyl group of 1-thyminyl and the π cloud of the neighboring nucleobase as well as Me(T)–Me(T) interactions plays the dominant role over other stacking interactions. This is because when $[\%T]_{DR-RD}$ are nearing zero, $[\Delta T_m]_{DR-RD}$ are also very close to zero (Fig. 1). This means that the other stacking interactions in the pairs **1–14** (Table 1) are relatively small.

(2) The relative ΔG°_{bp} decreases as the number of excess Ts in DR over RD increases (Fig. 3A).

(3) As the number of excess T increases in DR compared to isosequential RD, the relative stacking stabilization in DR increases because of the methyl-T effect [Me(T)–Me(T) + Me(T)–π(N), N = A/G/C/T, interactions] (Fig. 3B and Table 1).

(4) Clearly, this work shows that such methyl-T effect in the stabilization of DD duplexes should also play a key role in the overall ΔG°_{37} of DD. This methyl-T effect can however be only distinguished by the present comparison of DR *versus* RD in which we can fully control the presence of T in the D strand when the R strand can be constituted of only U. This means that using our present strategy, we compare only stabilization owing to the intrastrand T stacking interactions in the D strand and interstrand A:T base-pair interactions compared to intrastrand U interaction and A:U base-pair.

(5) The highest bonding energy is found for the single hydrogen bond, when the two negative atoms in the best donor and acceptor can freely adopt a collinear orientation. Clearly the pK_a matching in those cases are fully applicable. The role of pK_a matching for those cases which form two or three hydrogen bonds, as in the base-pairs of DNA–RNA duplexes which deviate from the collinear orientation, has not so far been energetically estimated. Electrostatic calculations in the gas phase however suggest that deviation from linearity can lead to a decrease in the hydrogen-bonding energy,[37] which is however difficult to compare with the solution state conformation. Since the N–H···O angle in base–base interactions in homo and hetero duplexes are known to vary by 3–23°,[36] it is therefore likely that our estimation of the relative strength of the base-pairing $[\Delta G^{\circ}_{bp}]_{DR-RD}$ from $\Sigma\Delta pK_{a[DR-RD]}$ values may have some intrinsic error, which we however can not estimate[14] accurately despite the fact that our pK_a measurement error is ±0.02, which means that we have an error of ±0.1 kJ mol^{-1} in our experimental $\Delta G^{\circ}_{pK_a}$ calculation. Our control study on the pK_a matching of ribo(G:C) and 2′-deoxyribo(G:C) showed a difference in $\Delta\Delta\Delta G^{\circ}_{pK_a}$ of only 1.0 kJ mol^{-1} for a base-pair,[14] owing to the fact that their respective sugar conformations represent two extreme conformations in the pentose-sugar pseudorotational cycle (which may bring about different energy cost for their respective hydration). This means that the A:U base-pairs are stabilized over the A:T base-pairs by at least 1.0 ± 0.2 kcal mol^{-1} by our pK_a matching procedure based on $\Delta\Delta\Delta G^{\circ}_{pK_a}$. Interestingly, this observation goes hand-in-hand with the recent comparative TROSY study[15] with homologous RNA and DNA sequences in the solution, which shows again 0.4 kcal mol^{-1} stabilization of A:U over the A:T base-pairs. Although recent high-level quantum-mechanical gas phase study of A:U *versus* A:T dimeric base-pairs suggests only a marginal difference (0.1 kcal mol^{-1}) in the stability of A:T and A:U base-pairs, this discrepancy can perhaps be attributed to their modeling in the gas phase compared to the thermodynamic parameters produced in the actual experimental solution studies,[14,15] as the authors[38] have rightly pointed out.

Further convincing experimental data that supports the fact that the A:U base-pairing is indeed stronger than A:T base-pairing can be found through the present work (Fig. 3A) that as the number of excess T increases in the D strand the relative strength of the base-pairing in the heteroduplexes simply decreases but the relative strength of the stacking forces increases (Fig. 3B) because of more number of Me–T interactions with the neighboring nucleobases.

Experimental

(A) pH-dependent ^{1}H NMR measurement

All NMR experiments were performed in Bruker DRX-500 and DRX-600 spectrometers. The NMR samples of all 2′-deoxy (**6a**–**10a**) and ribo (**6b**–**10b**) pairs of nucleoside 3′,5′-bis-ethylphosphates, Etp(d/rN)pEt, and 2′-deoxy (**1a**–**5a**) and ribo (**1b**–**5b**) pairs of nucleoside 3′-ethyl-phosphates were prepared in D_2O solution[14](concentration of 1 mM in order to rule out any chemical shift change owing to self-association) with

$\delta_{DSS} = 0.015$ ppm as internal standard. All pH-dependent NMR measurements have been performed at 298 K. The pH values [with the correction of deuterium effect] correspond to the reading on a pH meter equipped with a calomel microelectrode (in order to measure the pH inside the NMR tube) calibrated with standard buffer solutions (in H_2O) of pH 4, 7 and 10. The pD of the sample has been adjusted by simple addition of micro liter volumes of NaOD solutions (0.5 M, 0.1 M and 0.01 M). The assignments (see Supporting Information of ref. 14) for all compounds have been performed on the basis of selective homo-(^{1}H) and heteronuclear (^{31}P) decoupling experiments. All ^{1}H spectra have been recorded using 128 K data points and 64 scans.

(B) The pH titration of aromatic protons

The pH titration studies [over the range of pH 1.8 < pH < 12.2, with an interval of pH 0.2–0.3,] for 2′-deoxy (**6a**–**10a**) and ribo (**6b**–**10b**) pairs of nucleoside 3′,5′-bis-ethylphosphates, Etp(d/rN)pEt, and four 2′-deoxy (**1a**–**5a**) and ribo (**1b**–**5b**) pairs of nucleoside 3′-ethylphosphates.[14] All pH titration studies consist of ~20–33 data points and the corresponding Hill plots for all compounds are given in the Supporting Information of ref 14 and the pK_as shown in Table 1 (ref 14) have been calculated from Hill plot analysis.

(C) pK_a determination

The pH-dependent [over the range of pH 1.8 < pH < 12.2, with an interval of pH 0.2–0.3] ^{1}H chemical shifts (δ, with error ±0.001 ppm) for all compounds (for 2′-deoxy series: **1a**–**5a** and **6a**–**10a** as well as ribo series: **1b**–**5b** and **6b**–**10b**)[14] show a sigmoidal behavior (Fig. 1).[14] The pK_a determination is based on the Hill plot analysis using equation: $pH = \log((1 - \alpha)/\alpha) + pK_a$, where α represents fraction of the protonated species. The value of α is calculated from the change of chemical shift relative to the deprotonated (D) state at a given pH ($\Delta_D = \delta_D - \delta_{obs.}$ for deprotonation, where δ_{obs} is the experimental chemical shift at a particular pH), divided by the total change in chemical shift between neutral (N) and deprotonated (D) state (Δ_T). So the Henderson–Hasselbach type equation can then be written as $pH = \log((\Delta_T - \Delta_D)/\Delta_D) + pK_a$. The p$K_a$ is calculated from the linear regression analysis of the Hill plot.[14]

(D) T_m and corresponding Gibb's free energies of hybrid duplexes

T_m data and corresponding Gibb's free energy of duplex formation of 14 isosequencial DNA–RNA hybrid duplexes have been taken from Lesnik *et al.*[4]

Acknowledgements

Generous financial support from the Swedish Natural Science Research Council (Vetenskapsrådet), the Stiftelsen för Strategisk Forskning and Philip Morris USA Inc is gratefully acknowledged. We also thank Dr P. Acharya for help in the preparation of the manuscript.

References

1 B. I. Kankia and L. A. Marky, *J. Phys. Chem. B*, 1999, **103**, 8759.
2 K. B. Hall and L. W. McLaughlin, *Biochemistry*, 1991, **30**, 10606.
3 H. Han and P. B. Dervan, *Proc. Natl. Acad. Sci. U. S. A.*, 1993, **90**, 3806.
4 E. A. Lesnik and S. M. Freier, *Biochemistry*, 1995, **34**, 10807.
5 C. Escude, J. C. Francois, J. S. Sun, G. Ott, M. Sprinzl, T. Garestier and C. Helene, *Nucleic Acids Res.*, 1993, **21**, 5547.
6 L. Ratmeyer, R. Vinayak, Y. Y. Zhong, G. Zon and W. D. Wilson, *Biochemistry*, 1994, **33**, 5298.
7 A. Kornberg and T. A. Baker, *DNA Replication*, W. H. Freeman, New York, 1992.
8 R. Okazaki, A. Sugino, T. Okazaki, Y. Imae, R. L. Kainuma-Kuroda, T. Ogawa, M. Arisawa and Y. Kurasowa, in *DNA Synthesis in vitro*, ed. R. Wells, R. B. Inman, University Park Press, Baltimore, 1973, pp. 83–106.
9 H. Varmus, *Science*, 1988, **240**, 1427.
10 *Prospects for Antisense Nucleic Acid Therapy of Cancer and AIDS*, ed. E. Wickstrom, 1991, p 269 pp.
11 S. T. Crooke, *Annu. Rev. Med.*, 2004, **55**, 61.
12 I. Lebedeva and C. A. Stein, *Annu. Rev. Pharmacol. Toxicol.*, 2001, **41**, 403.
13 J. B. Opalinska and A. M. Gewirtz, *Nat. Rev. Drug Discovery*, 2002, **1**, 503.
14 P. Acharya, P. Cheruku, S. Chatterjee, S. Acharya and J. Chattopadhyaya, *J. Am. Chem. Soc.*, 2004, **126**, 2862.
15 I. Vakonakis and A. C. LiWang, *J. Am. Chem. Soc.*, 2004, **126**, 5688.
16 K. Snoussi and J. L. Leroy, *Biochemistry*, 2001, **40**, 8898.
17 W. G. Scott, J. B. Murray, J. R. Arnold, B. L. Stoddard and A. Klug, *Science*, 1996, **274**, 2065.
18 Y. Pan and A. D. MacKerell, Jr., *Nucleic Acids Res.*, 2003, **31**, 7131.
19 S. Wang and E. T. Kool, *Biochemistry*, 1995, **34**, 4125.
20 Y. Umezawa and M. Nishio, *Bioorg. Med. Chem.*, 2000, **8**, 2643.
21 Y. Umezawa and M. Nishio, *Nucleic Acids Res.*, 2002, **30**, 2183.
22 M. A. Viswamitra, O. Kennard, P. G. Jones, G. M. Sheldrick, S. Salisbury, L. Favello and Z. Shakked, *Nature*, 1978, **273**, 687.
23 S. M. Freier and K. H. Altmann, *Nucleic Acids Res.*, 1997, **25**, 4429.
24 C. A. Hunter, *J. Mol. Biol.*, 1993, **230**, 1025.
25 A CH(σ)–π interaction[18–20,34] is said to be a weak attractive hyperconjugative donor–acceptor interaction between a soft acid and a soft base. This interaction mainly involves dispersion, charge transfer and electrostatic forces, in which the sigma orbital of the CH bond preferably chooses the site of H to be interacted with the πsurface of a soft base. This cooperative interaction is orientation dependent, additive in enthalpy, entropically favored, effective in water as well as in nonpolar media, involving around 1 kcal mol^{-1} in energy. Kool and coworkers established that CH(σ)–π interaction[18–20,34] together with Me–Me interactions are stabilizing for the duplex. We here show that both CH(σ)–π interaction and Me–Me interaction,[18,33] in the dissected form, as shown in this work, also act as the stabilizing force for the duplex, except for the fact that the former is three times stronger than the latter (unlike the weighting of 57:43 for CH–π and Me–Me effects by methyl group by Kool *et al.*). The CH_3(T)–π interaction (when Ts are located in the middle of the DNA duplex) has been shown to be playing towards both 5′ends of the individual single strands, which forms the DNA double helix. The methyl group (compare uracil moiety with 5-methyluracil enhances[18,28] the base stacking proficiency are evident from the T_m study of methylated and non-methylated oligos) is stabilizing the duplex. The role of any hydrophobic effects for the methyl group has also been ruled out by the fact that the methylation actually increases the water solubility of uracil.[35] The Me–Me interaction has been attributed by Kool *et al.* by invoking that the polarizability of the methyl group leading to favorable van der Waal interaction, and in case of two neighboring methyl groups the dipole-induced dipole interaction has been said to be responsible as the stabilizing force.
26 P. J. Hagerman, *Nature*, 1986, **321**, 449.
27 K. Umemoto, M. H. Sarma, G. Gupta and R. H. Sarma, *Biochemistry*, 1990, **29**, 4714.
28 S. O. Shan and D. Herschlag, *Proc. Natl. Acad. Sci. U. S. A.*, 1996, **93**, 14474.
29 S. O. Shan, S. Loh and D. Herschlag, *Science*, 1996, **272**, 97.
30 S. Acharya, J. Barman, P. Cheruku, S. Chatterjee, P. Acharya, J. Isaksson and J. Chattopadhyaya, *J. Am. Chem. Soc.*, 2004, **126**, 8674.
31 J. W. Larson and T. B. McMahon, *J. Am. Chem. Soc.*, 1983, **105**, 2944.
32 J. Chen, M. A. McAllister, J. K. Lee and K. N. Houk, *J. Org. Chem.*, 1998, **63**, 4611.
33 M. F. Grenier-Loustalot, J. Bonastre and P. Grenier, *J. Mol. Struct.*, 1980, **65**, 249.
34 F. A. Jalbaout, A. E. Basso, M. R. Pontes and D. Das, *J. Mol. Struct. (THEOCHEM)*, 2004, **677**, 167.
35 L. C. Sowers, B. R. Shaw and W. D. Sedwick, *Biochem. Biophys. Res. Commun.*, 1987, **148**, 790.
36 W. Saenger, *Principles of Nucleic Acid Structure*, Springer Verlag, Berlin, 1988.
37 C. Pimentel and A. L. McClellan, *The Hydrogen Bond*, Freeman, San Francisco, CA, 1960.
38 A. Perez, J. Sponer, P. Jurecka, P. Hobza, F. J. Luque and M. Orozco, *Chem. Eur. J.*, 2005, **11**, 5062.