Published online 15 January 2010

Nucleic Acids Research, 2010, Vol. 38, No. 9 ***3133–3146***
doi:10.1093/nar/gkp1237

Aromatic N versus aromatic F: bioisosterism discovered in RNA base pairing interactions leads to a novel class of universal base analogs

Alrun N. Koller[1], Jelena Božilović[2], Joachim W. Engels[2] and Holger Gohlke[1,*]

[1]Department of Mathematics and Natural Sciences, Institute of Pharmaceutical and Medicinal Chemistry, Heinrich-Heine-University, 40225 Düsseldorf and [2]Department of Biochemistry, Chemistry and Pharmacy, Institute for Organic Chemistry, Goethe-University, 60438 Frankfurt am Main, Germany

Received November 12, 2009; Revised and Accepted December 23, 2009

ABSTRACT

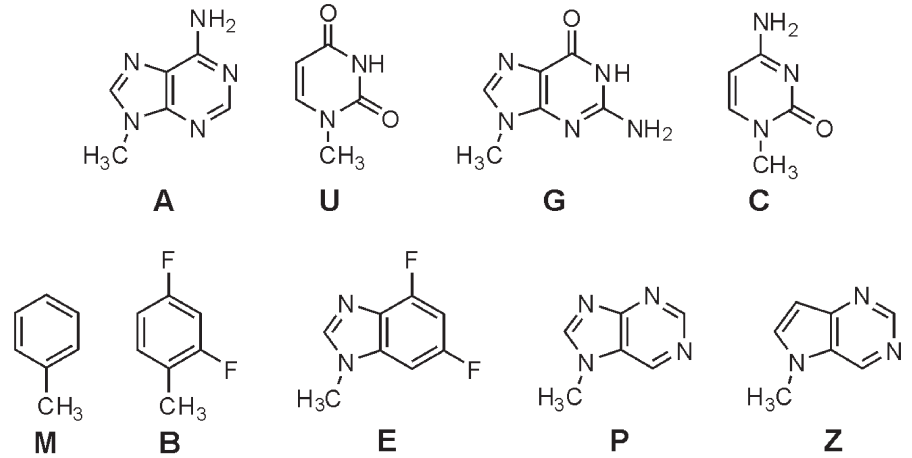
The thermodynamics of base pairing is of fundamental importance. Fluorinated base analogs are valuable tools for investigating pairing interactions. To understand the influence of direct base–base interactions in relation to the role of water, pairing free energies between natural nucleobases and fluorinated analogs are estimated by potential of mean force calculations. Compared to pairing of AU and GC, pairing involving fluorinated analogs is unfavorable by 0.5–1.0 kcal mol^{-1}. Decomposing the pairing free energies into enthalpic and entropic contributions reveals fundamental differences for Watson–Crick pairs compared to pairs involving fluorinated analogs. These differences originate from direct base–base interactions and contributions of water. Pairing free energies of fluorinated base analogs with natural bases are less unfavorable by 0.5–1.0 kcal mol^{-1} compared to non-fluorinated analogs. This is attributed to stabilizing C–F$\cdots$H–N dipolar interactions and stronger N$\cdots$H–C hydrogen bonds, demonstrating direct and indirect influences of fluorine. 7-methyl-7*H*-purine and its 9-deaza analog (Z) have been suggested as members of a new class of non-fluorinated base analogs. Z is found to be the least destabilizing universal base in the context of RNA known to date. This is the first experimental evidence for nitrogen-containing heterocylces as bioisosteres of aromatic rings bearing fluorine atoms.

INTRODUCTION

Hydrogen bonds, base stacking, and solvation are the three predominant forces governing the stability of nucleic acid structures. To probe these interactions, a common approach is to replace natural bases with analogs in which functional groups are added, deleted, blocked, or rearranged. Size and shape of the analogs are preserved as close as possible to natural bases. These 'non-polar nucleoside isosteres' (NNIs) (1) allow detecting predominant forces within nucleic acid structures without introducing steric effects. NNIs have also proven to be valuable tools in biochemical experiments (2). Prominent classes of NNIs are fluorinated indoles and benzimidazoles as well as fluorinated toluenes (1). When paired against natural bases, the fluorinated analogs destabilize DNA and RNA duplexes. However, they exhibit little binding sequence specificity and have the nature of universal bases (3,4). These observations imply a lack of Watson–Crick base pairing involving hydrogen bonds to fluorine (3,5) in agreement with findings that covalently bound fluorine hardly ever acts as acceptor for available Bronsted acidic sites in the presence of competing heteroatom acceptors (6–9). Rather, the strong electronegativity of the fluorine atom and, thus, the strong dipole of the C–F bond favors dipole–dipole interactions in molecular recognition (4,10–18).

Due to its fundamental importance for biology and supramolecular chemistry, the thermodynamics of base pairing have been studied for many years (19). In water, the biologically relevant solvent, pairing energies are weaker than stacking interactions, and no direct measurements of pairing can be made (20). Instead, interactions between strands of well-defined DNA and RNA duplexes

*To whom correspondence should be addressed. Tel: +49 211 81 13662; Fax: +49 211 81 13847; Email: gohlke@uni-duesseldorf.de
Correspondence may also be addressed to Joachim W. Engels. Tel: +49 69 798 29150; Fax: +49 69 798 29148; Email: joachim.engels@chemie.uni-frankfurt.de



Scheme 1. Structures of natural RNA bases and base analogs. Key: **U** = 1-methyl-uracil; **C** = 1-metyhl-cytosine; **A** = 9-methyl-adenine; **G** = 9-methyl-guanine; **M** = toluene; **B** = 2,4-difluorotoluene; **E** = 4,6-difluoro-1-methyl-1*H*-benzimidazole; **P** = 7-methyl-7*H*-purine and **Z** = 5-methyl-5*H*-pyrrolo[3,2-d]pyrimidine.

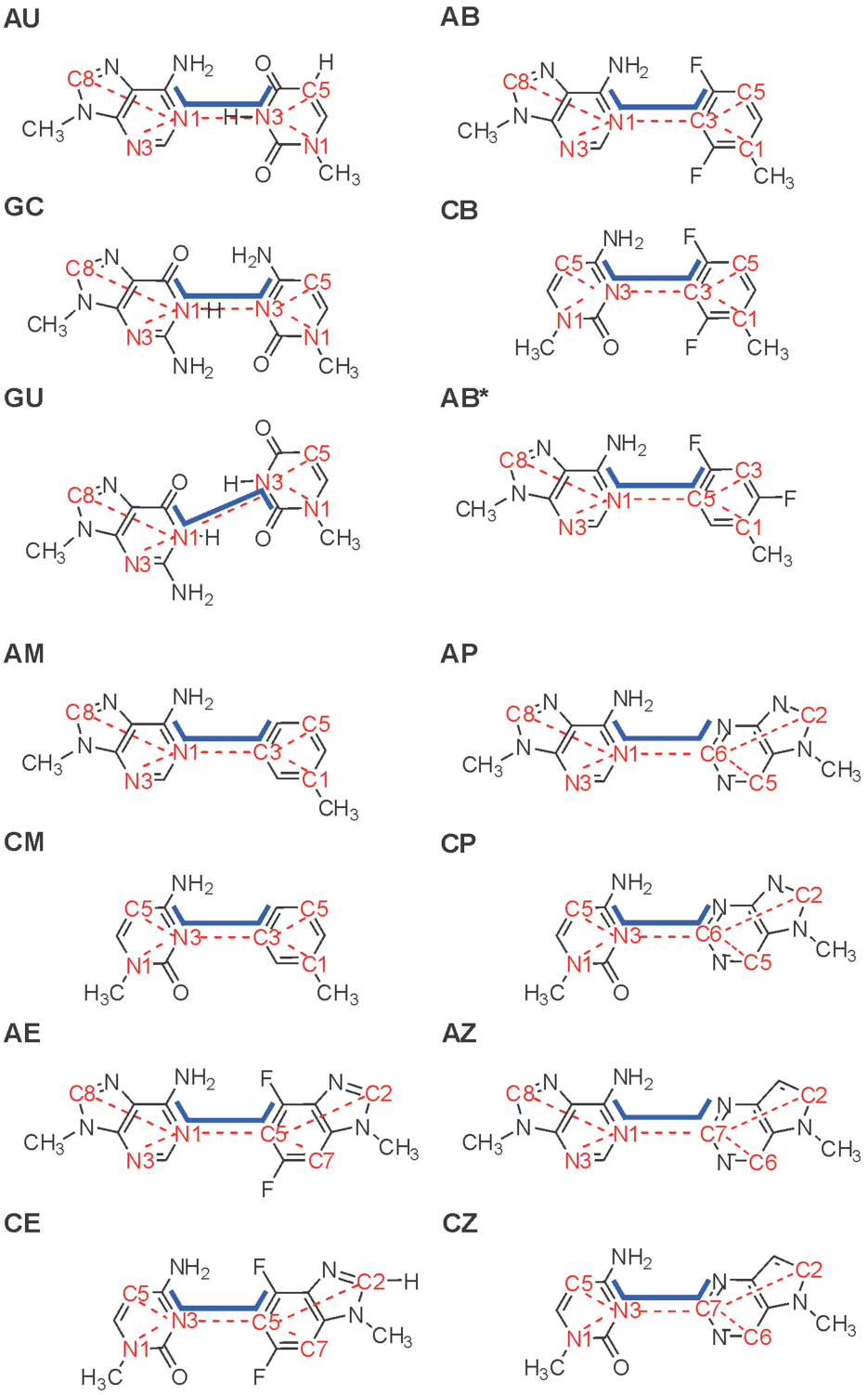
have been studied. For pairing involving natural nucleobases, this has led to an estimate of a free-energy increment associated with a single hydrogen bond in the range of 0.8–2.2 kcal mol^{-1} (21–24). However, interactions between strands also involve base stacking and factors linked to conformational changes in the backbone (20,25), and stacking and pairing interactions may even be coupled (26). Such difficulties can be overcome if base pairing is investigated by computational approaches. For pairing involving natural bases, recent free-energy estimates based on improved simulation methods and force fields are in good agreement with experimental values (20,27). For pairing between adenine (**A,** Scheme 1) and a fluorinated base analog, 2,4-difluorotoluene (**B**), DFT calculations have revealed a total interaction energy of $\sim$−3 kcal mol^{-1}, a value equivalent to a weak hydrogen bond (28). Close values of −3.8 and −4.2 kcal mol^{-1} have been determined by quantum mechanical calculations at the MP2 level (29,30). When related to a 1 M gas-phase reference state, however, **AB** pairing interactions have been found to be disfavorable (28). Regarding free-energy calculations, **AB** interactions so far have only been investigated in the context of duplex DNA, giving rise, again, to the difficulties mentioned above (31,32).

Based on our previous studies that incorporated fluorinated analogs as universal bases into 12-mer RNA duplexes (3,33–36) and a study of fluorobenzene self-pairing in duplex RNA (18), we set out to investigate pairing interactions between natural bases and fluorinated and non-fluorinated base analogs in aqueous solution. For this, we performed potential of mean force (PMF) calculations between natural and (non-)fluorinated bases in a Watson–Crick-like orientation in comparison to conventional **AU** and **GC** pairs. The PMF is a fundamental measure of the interactions between solutes in solution and, in our case, gives a pairing free energy as a function of the separation distance between bases. These calculations provide answers that help in understanding the influence of direct base–base interactions in relation to the role of water in the process. In addition, they demonstrate the prominent role of fluorine in base pairing. Considering that nitrogen-containing heterocycles have been proposed as mimics of aromatic rings bearing fluorine atoms (37), we then predicted and experimentally verified a 7-*N*-linked purine (**P**) and a 9-deaza analog **(Z**) as members of a new class of non-fluorinated base analogs. In the context of RNA duplexes, **Z** is the least destabilizing universal base known to date. To the best of our knowledge, this is the first experimental evidence of nitrogen-containing heterocylces as bioisosteres of aromatic rings bearing fluorine atoms.

MATERIALS AND METHODS

Base pair configurations

Natural bases and base analogs investigated in this study are shown in Scheme 1. Structures of the natural RNA bases **A**, **C**, **G** and **U** were taken from the nucgen module of AMBER8 and modified by replacing the ribose moiety with a methyl group. Similarly, natural purine/pyrimidine bases were modified to toluene (**M**), 2,4-difluorotoluene (**B**), 4,6-difluoro-1-methyl-1*H*-benzimidazole (**E**), 7-methyl-7*H*-purine (**P**) and 5-methyl-5*H*-pyrrolo[3,2-d]pyrimidine (**Z**). Atomic charges of the base analogs were determined following the RESP methodology (38). In the case of the natural bases, the atomic charges of the base atoms were taken from the AMBER library (39). The charges of the methyl group atoms at the glycosidic site were calculated with the RESP methodology (38) such that the net charge of the modified natural base is zero. The same level of theory was used to calculate molecular dipoles. Bond dipoles were calculated from atomic partial charges and the equilibrium bond distance. The base analogs were then paired with **A** and **C**, respectively. For obtaining configurations of the contact pairs, the respective bases were superimposed onto base pairs of a canonical A-form duplex RNA. Configurations of Watson–Crick base



Scheme 2. Base pair configurations for which PMFs were calculated. Key: angle restraints used to maintain the alignment of the base pairs are marked with dashed red lines. A dihedral angle restraint (blue) was used to prevent the bases from twisting. The distance for the PMF calculations was measured between the inner two atoms of the dihedral angle restraint, respectively.

pairs **AU** and **GC** were taken from a canonical A-form duplex RNA. The configuration of the wobble base **GU** was taken from an experimental RNA structure (40) (Nucleic Acid Database code: AR0009). All investigated configurations are depicted in Scheme 2.

Umbrella sampling simulations

Umbrella sampling simulations (41) were performed with the AMBER8 molecular simulation package (42). In an umbrella sampling simulation, the system is restrained to a narrow range of the conformational space by applying

a quadratic biasing potential $V_i(r)$. Here, the distance r between the two inner atoms of the dihedral angle restraint (Scheme 2) was chosen as a reaction coordinate. The parm94 force field (39), which has a very good balance of intermolecular interaction terms for nucleobases (43–45), was used for all simulations. The base pairs were placed in an octahedron of TIP3P (46) water molecules with a minimum distance of 15 Å between the base pair atoms and the surface of the octahedron and an initial distance of $r = 3$ Å. All simulations were carried out in the NVT ensemble at 300 K under periodic boundary conditions using a time step of 2 fs and the Berendsen thermostat (47). The SHAKE procedure (48) was applied to constrain all bonds involving hydrogens with a tolerance of 10^{-6} Å. A cutoff of 8.0 Å was used for non-bonded interactions, and long-range electrostatic interactions were treated with the particle mesh Ewald (PME) method (49). No influence of box size or cut off on the free-energy curves was observed (Supplementary Figures S1 and S2).

All base pair configurations were initially energy minimized for 500 cycles using the steepest descent and the conjugated gradient method with initial position restraints on the molecules. Over 50 ps, the systems were heated from 100 to 300 K at constant volume, solute atoms were still restrained. The solvent density was equilibrated by simulating the systems for 50 ps in the NPT ensemble at 1 atm and 300 K, still with harmonic restraints on the solute atoms. In the last equilibration step, a NVT simulation was conducted over 100 ps with a stepwise reduction of the applied restraints to zero and a subsequent unrestrained simulation of 200 ps. The equilibrated system was then used as the starting configuration for the first step in the umbrella sampling simulations.

The separation distance r was sampled from 3 to 10 Å in 1 Å steps with an applied restraining force of 1–10 kcal mol^{-1} Å^{-2}. The equilibrated base pair configuration was used as starting structure for the first step ($r = 3$ Å). The last configuration of the first step was then used as starting structure for the next step and so forth. For each step, an equilibration phase of 0.2 ns and a production phase of 0.4 ns was used. Each step was sampled at least twice.

To avoid the bases to move rapidly towards a stacked configuration during the simulations, the bases were kept aligned by angle and torsional restraints. The angle restraints were defined according to Stofer *et al.* (20) (red dashed lines in Scheme 2). Target values of the angle restraints are given in Supplementary Table S1. Additionally, the propeller angle between the two bases was restrained to zero degree in order to enforce a planar configuration of the bases (blue lines in Scheme 2). To investigate the influence of restraining the propeller angle on the free-energy curves, PMFs for the Watson–Crick base pair **GC** were computed with and without such restraint (Supplementary Figure S3). Restraining the propeller angle yields a pairing free energy that is ~0.9 kcal mol^{-1} more favorable compared to no torsional restraint. This can be explained as follows: for a contact base pair, the hydrogen bonding interactions already restrict the configurational space accessible to the bases (20). In contrast, no such restriction exists for separated bases. Hence, the effect of the propeller angle restraint is larger for the separated bases, and the state of separated bases is disfavored compared to the contact pair.

We expect that this effect will be smaller if pairs between natural bases and base analogs are considered. This is because interactions between these bases in a contact pair are less geometrically restricting. Hence, the states of separated bases and bases in a contact pair are more similar then, whether the propeller angle is restrained or not.

PMF calculations

To determine the PMF $W_i(r)$ from the biased frequency distributions $P_i(r)$ obtained by the umbrella sampling simulations performed for each step i, the $P_i(r)$ are post-processed with the weighted histogram analysis method (WHAM) [50]:

$$W_i(r) = -k_B T \ln P_i(r) - V_i(r) + C_i, \quad (1)$$

where k_B is the Boltzmann constant, T the temperature of the system, and $V_i(r)$ the restraining potential used for step i. The histograms $P_i(r)$ are combined and the factors C_i are iteratively fit to yield an optimal solution that satisfies the continuity constraint. Finally, the resulting PMF is shifted vertically such that the value at the longest distance sampled becomes zero. Statistical uncertainties for each point of the PMF were calculated with a Monte Carlo bootstrap error analysis (51) as implemented within the WHAM program using 100 fake data sets.

Free-energy decomposition

The entropic component $-T\Delta S$ of the free-energy curves was computed according to MacCallum *et al.* (52) using a finite-difference approximation [Equation (2)].

$$T\Delta S_{(T)} = -T\frac{\Delta G(T+\Delta T) - \Delta G(T-\Delta T)}{2\Delta T}. \quad (2)$$

The enthalpic component ΔH is then obtained as

$$\Delta H_{(T)} = \Delta G_{(T)} + T\Delta S_{(T)}. \quad (3)$$

Additional simulations at 285 and 315 K were performed for base pairs **AU** and **AB**, and all PMFs were shifted to zero at the largest separation distance investigated prior to applying Equations (2) and (3).

RESULTS AND DISCUSSION

Pairing free energies of natural RNA bases

Initially, the PMFs of Watson–Crick RNA base pairs **AU** and **GC**, and the wobble base pair **GU** were calculated (Scheme 2). The free-energy curves for base separation are shown in Figure 1a. The curves are all adjusted to zero for the maximal separation between the bases of 10.0 Å. The statistical uncertainty of the free-energy values were estimated by a Monte Carlo bootstrap

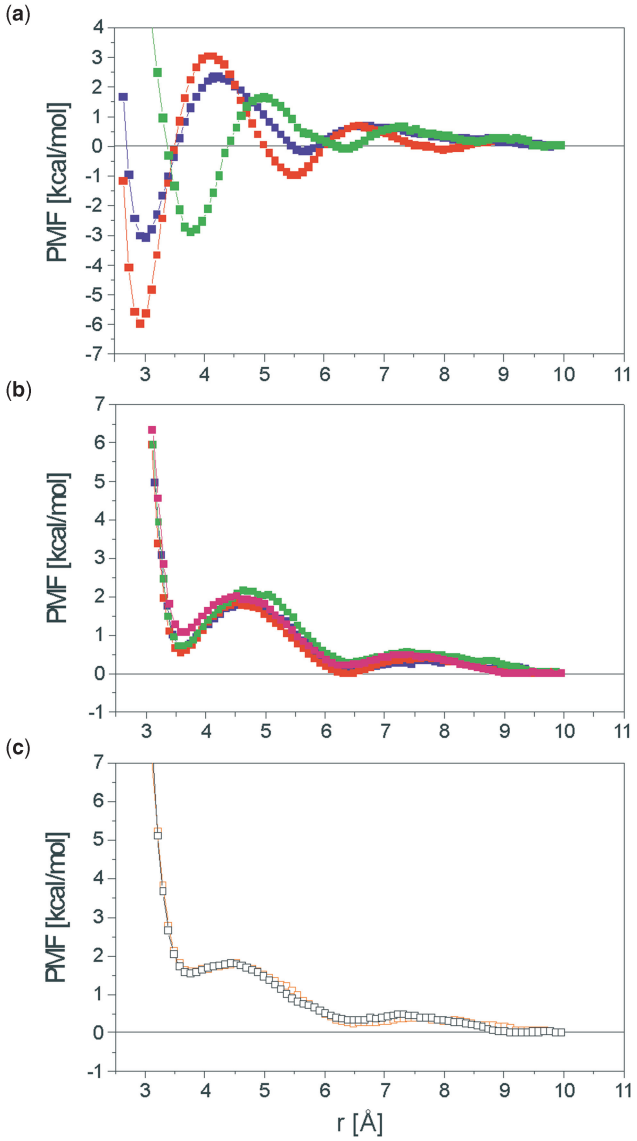


Figure 1. Free-energy profiles as a function of separation distance for (**a**) the natural base pairs **AU** (blue) and **GC** (red) and the wobble base pair **GU** (green); (**b**) the base pairs **AB** (blue), **CB** (red), **AE** (green) and **CE** (magenta); (**c**) the base pairs **AM** (black) and **CM** (orange).

analysis (51) and is <0.1 kcal mol^{-1}. The distance between N1 of the purine and N3 of the pyrimidine systems was chosen as the reaction coordinate in all cases. This leads to an offset of ~0.8 Å for **GU** if the distance between atoms involved in hydrogen bonding is considered instead. Taking this into account, the three base pairs show the same behavior. With decreasing separation, a first energy barrier is encountered at ~6.5 Å, followed by a shallow minimum at ~5.6 Å (Table 1), representing a water-separated base pair configuration as it has been demonstrated by Stofer *et al.* (20) for DNA base pairs. This minimum is separated by a second energy barrier at ~4.2 Å, when the bridging water molecules are expelled from between the bases. Finally, the contact minimum at the hydrogen bonding distance of ~3.0 Å is reached, representing the pairing free energy. The positions of the energy minima and barriers are in excellent agreement with those found by Stofer *et al.* (20) for **AT** and **GC** base pairs.

The **GU** base pair has a comparable thermodynamic stability to that of **AU** and is nearly isomorphic (53). This is also reflected in our results: energy minima found for **GU** are only slightly less favorable than those of **AU** (Table 1)**.** The positions of the energy minima and barriers of **GC** occur at distances 0.1 Å shorter than for **AU** due to the stronger interactions between the **GC** bases. This is also reflected in lower free energies of the minima in the **GC** case. The value of −5.98 kcal mol^{-1} at the contact minimum agrees well with the pairing free energy of −5.7 kcal mol^{-1} found by Stofer *et al.* (20). Our results furthermore agree with experimental studies that estimated the contribution of a single hydrogen bond to the pairing free energy as 1.7–2.2 kcal mol^{-1} (21–24). Together, with the statistical quality of our estimates, these findings give us confidence in the simulation method and the applied force field, which has been shown to have a very good balance of intermolecular interaction terms for nucleobases (43–45).

Pairing free energies of fluorinated base analogs

Next, we investigated the pairing between fluorinated base analogs and natural RNA bases in order to elucidate differences in direct base–base interactions and interactions with the solvent that give rise to the behavior of the base

Table 1. Thermodynamic properties of the investigated base-pair configurations[a]

	AX							GX	CX					
X:	U	B	B*	E	M	P	Z	U	G	B	E	M	P	Z
M1	3.02	3.65	3.67	3.63	3.77	3.67	3.67	3.77	2.92	3.58	3.67	3.77	3.77	3.77
B	4.23	4.75	4.80	4.63	4.52	4.23	4.42	4.98	4.05	4.52	4.52	4.42	4.33	4.33
M2	5.64	6.45	6.48	6.46	6.48	6.48	6.48	6.39	5.45	6.39	6.30	6.39	6.30	6.39
ΔG_{M1}	−3.07	0.70	1.02	0.71	1.60	0.83	0.78	−2.91	−5.98	0.55	1.07	1.55	1.12	0.99
ΔG_{B}	2.33	1.89	2.02	2.19	1.80	0.97	0.96	1.61	3.05	1.83	2.01	1.80	1.25	1.10
ΔG_{M2}	−0.18	0.15	0.30	0.29	0.26	0.24	0.21	−0.12	−0.99	−0.01	0.20	0.32	0.23	0.12

[a]M1, B and M2 (in Å) are, respectively, the distance of the inner two atoms of the dihedral angle restraint (see Scheme 2) at the contact minimum, the desolvation barrier and the water-separated minimum. ΔG_{M1}, ΔG_{B} and ΔG_{M2} (in kcal mol^{-1}) are, respectively, the free energies of the contact minimum, the desolvation barrier, and the water-separated minimum. The statistical uncertainty of the free energy values were estimated by a Monte Carlo bootstrap analysis (51) and is <0.1 kcal mol^{-1} in all cases.

Table 2. Thermodynamic properties of 12-mer duplex RNAs (5′-CUU UUC XUU CUU-3′ paired with 3′-GAA AAG YAA GAA-5′)[a]

	Y = A		Y = C		Y = G		Y = U	
X	T_m	ΔG	T_m	ΔG	T_m	ΔG	T_m	ΔG
U[b]	37.8	11.9	30.4	9.8	38.6	11.9	30.1	9.7
M[b]	23.0	7.7	22.6	7.6	23.5	7.9	23.1	7.7
B[b]	27.4	9.0	27.3	8.9	27.6	9.0	27.9	9.1
E[b]	28.4	9.2	28.7	9.2	29.4	9.5	29.3	9.5
P[c]	34.5	9.9	32.2	9.7	35.1	10.3	30.8	9.1
Z[c]	35.8	10.1	35.0	10.4	35.4	10.1	33.1	9.9

[a]All measurements were performed in a phosphate buffer (140 mM NaCl, 10 mM Na_2HPO_4, 10 mM NaH_2PO_4); T_m: in °C; ΔG: in kcal mol^{-1} at $T = 298$ K.
[b]Values taken from Parsch *et al.* (3); experimental uncertainties: T_m: ±0.2°C; ΔG: ±2%.
[c]Present work. Experimental uncertainties: T_m: ±0.2°C; ΔG: ±2%.

analogs as universal bases. The calculated free-energy curves for the fluorinated bases **B** and **E** paired with **A** and **C**, respectively, are shown in Figure 1b. They are adjusted to zero for the maximal separation between the bases. The free-energy curves of the four pairs are very similar in shape. When compared to the curves for the natural RNA base pairs, the curves are qualitatively similar, too, although differences occur with respect to the positions and free-energy values of the minima and barriers. Upon lowering the distance between the bases, a first energy barrier occurs at ~7.6 Å, followed by a water-separated configuration with an energy minimum at ~6.4 Å. This minimum is separated by a barrier at ~4.7 Å from the contact minimum at ~3.7 Å. Compared to the Watson–Crick base pairs, the distances between the bases at the contact minimum are larger by ~0.7 Å.

The free energies of base pairing involving fluorinated base analogs are given in Table 1. The differences in the base pairing free energies of **B** and **E** with either one of the natural bases are small (0.15 and 0.36 kcal mol^{-1}), which supports the universal base pairing properties of **B** and **E** observed in our previous experiments (3). The base pairing involving fluorinated base analogs is unfavorable as demonstrated by pairing free energies of 0.55 to 1.07 kcal mol^{-1}, in stark contrast to Watson–Crick RNA base pair formation of natural bases. The results are consistent with the destabilizing effect observed on 12-mer RNA if a natural base is replaced by a fluorinated analog (Table 2) (3).

Decomposition of the pairing free energy into enthalpic and entropic components

To gain further insight into the forces that govern pairing between natural bases and fluorinated base analogs, we decomposed the PMFs of base pairs **AU** and **AB** into enthalpic and entropic contributions. For this, PMFs were calculated at three different temperatures, 285, 300 and 315 K, and the entropic contribution ($T\Delta S$) was determined by numerical differentiation of the resulting curves [Equation (2)]. The enthalpic component (ΔH) was then

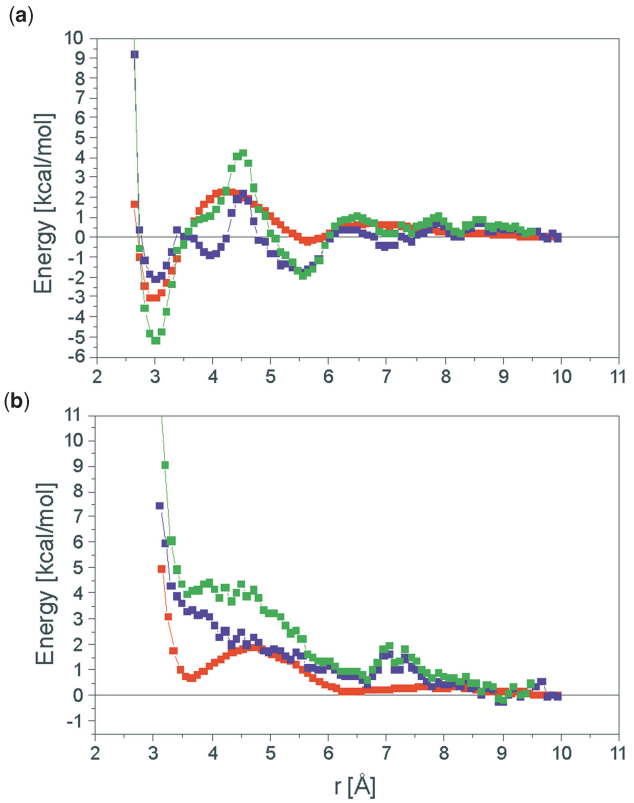


Figure 2. Enthalpic (ΔH; green) and entropic ($T\Delta S$; blue) components of the pairing free energies as a function of the separation distance for the base pairs (**a**) **AU** and (**b**) **AB**. In addition, the free-energy profiles at 300 K are given in red.

obtained as the sum of the free-energy curve and the entropic component [Equation (3)].

The decomposition results are shown in Figure 2, and the free-energy curves of the base pairs at different temperatures are shown in Supplementary Figure S4. ΔH and $T\Delta S$ contributions at the barrier ~4.5 Å are very similar for both the Watson–Crick complex and the one involving the fluorinated base analogs. Expelling water from in between the bases is enthalpically disfavored by ~4 kcal mol^{-1}, but entropically favored by ~2 kcal mol^{-1}. This results in a free energy cost of ~2 kcal mol^{-1}. Upon formation of the contact pairs, however, the enthalpic and entropic components show fundamental differences between **AU** and **AB**. For the Watson–Crick base pair **AU,** base pairing is enthalpy driven by ~5 kcal mol^{-1}, as expected by the formation of hydrogen bonds between the bases. At the same time, the geometrically restricted hydrogen bonds limit the mutual configurational freedom of the base pair, resulting in an overall entropically disfavored (~2 kcal mol^{-1}) base pairing.

In contrast, for **AB**, base pairing is enthalpically disfavored by ~4 kcal mol^{-1}. Apparently, the loss of the base–water interaction energy is not overcompensated by the formation of base–base interactions, as demonstrated by a very similar enthalpic component at the location of the barrier and the contact minimum. Schweitzer *et al.* (1) already proposed a net energetic repulsion between

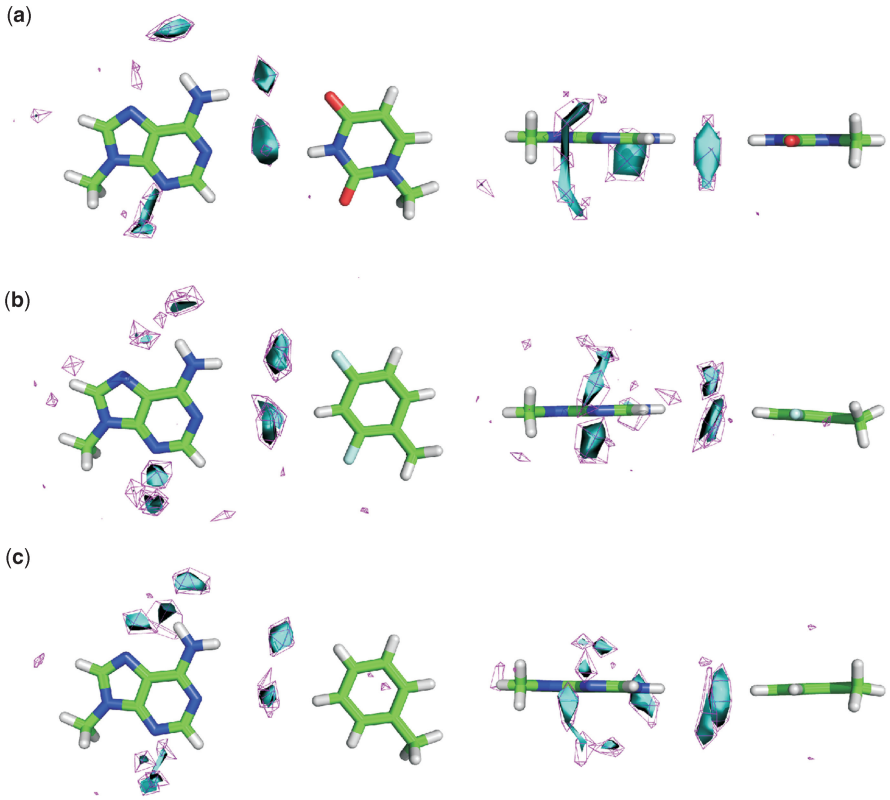


Figure 3. Top (left panel) and side view (right panel) of the water density in the umbrella sampling simulation at the water-separated minimum for base pairs (**a**) **AU**, (**b**) **AB**, (**c**) **AM,** respectively. Cyan solid regions represent a 2.5-fold higher and magenta mesh regions a 2.0-fold higher water density than the bulk density of 0.99 g cm^{-3}.

A and **B** and suggested the cost of removing waters of solvation bound to **A** as a reason (54,55). In fact, computed water densities for **AU** and **AB** configurations at the water-separated minima show two solvating waters between the bases, which are expelled upon formation of the contact pair (Figure 3). In contrast to the **AU** case, little energetic compensation for breaking these hydrogen bonds is available in the **AB** case due to **B**'s inability to form measurable hydrogen bonds with **A** (3,56). This is corroborated by quantum chemical studies, which predict a significantly lower stability for the pairing of fluorinated base analogs and **A** compared to a Watson–Crick pair (29,57,58).

The enthalpic component to base pair formation is partly compensated, however, by an entropic component of ~3.3 kcal mol^{-1} at the contact minimum. This contribution is 1.3 kcal mol^{-1} higher than the entropic component at the barrier. While the origin of this gain is not clear, the value demonstrates that an **AB** base pair is not as strongly locked in as an **AU** base pair. This is in agreement with MD simulations of DNA and RNA containing **AB** base pairs that have found an increased mobility and breathing frequency at base analog sites (28,59).

Role of fluorine in base pairing

Although replacement of the natural base **U** in the Watson–Crick pair **AU** by **B** destabilizes the RNA 12-mer by 2.9 kcal mol^{-1}, inserting **M** as a non-fluorinated analog leads to a larger destabilization of 4.2 kcal mol^{-1} (Table 2) (3). Thus, we performed PMF calculations for **M** paired with **A** and **C**, respectively, in order to determine the role of fluorine in base pairing.

The resulting free-energy profiles of **AM** and **CM** are shown in Figure 1c. In agreement with our experiments (Table 2) (3), no differences in the free-energy profile with respect to the pairing base are observed, which supports the universal base character of **M**. The location of the water-separated and contact minima and the barrier in between them is almost identical for base pairs involving **B** and **M**. All curves also agree with respect to the free-energy values at the water-separated minimum (~0.2 kcal mol^{-1}) and the barrier (~1.9 kcal mol^{-1}) (Table 1). Apparently, interactions between the bases and the solvent are similar irrespective of using a fluorinated or non-fluorinated base analog as one of the pairing partners. Likewise, a similar free-energy cost is required for expelling the waters upon forming the contact pair. In agreement with these data, computed water densities for **AB** and **AM** configurations at the water-separated minima both show two waters located between the bases (Figure 3). This indicates that the base–water interactions are mostly determined by the natural RNA base, which forms hydrogen bonds with water. In contrast, interactions occurring in the ring plane between either **B** or **M** and water are weak

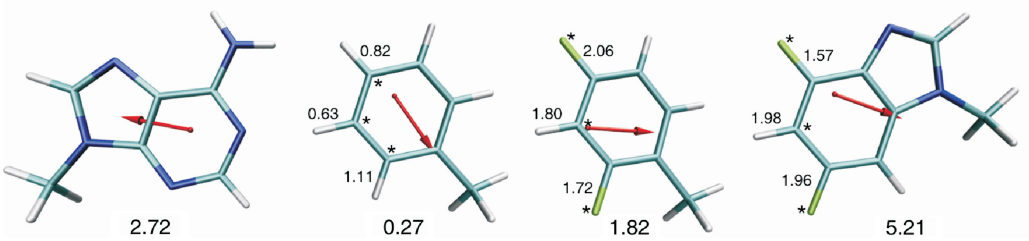


Figure 4. Structures of **A**, **M**, **B** and **E**, with Watson–Crick-like orientations of **A** and the base analogs, respectively. The red arrows depict the orientation of the molecular dipole moments. The magnitude of the molecular dipole moment is given below each structure. Numbers next to C–F or C–H bonds indicate the magnitude of bond dipoles. The negatively polarized ends of the bond dipoles are marked by asterisks, respectively. All dipoles are given in Debye.

and not influenced by the base type. In summary, we conclude that stability differences observed for fluorinated or non-fluorinated base analogs in the context of duplex RNA do not arise from differences in the interactions to water located in the ring plane of these base analogs.

The free-energy curves of **AM** and **AB** (**AE**) differ markedly at the contact minimum. Whereas the energy difference between the minimum and the barrier is only 0.20 kcal mol^{-1} in the case of **AM**, the free energy decreases by 1.19 kcal mol^{-1} (1.49 kcal mol^{-1}) when the fluorinated base analog **B** (**E**) contact **A**. A similar decrease is seen if the pairing base is **C** instead of **A**. Apparently, while the overall pairing free energy for natural base/fluorinated base analog complexes is unfavorable, it is so by only 0.55–1.07 kcal mol^{-1}. On the contrary, pairing of **M** with natural bases is unfavorable by 1.55–1.60 kcal mol^{-1}. Notably, these results are in agreement with differences in the stabilities of RNA duplexes containing either **M** or **B** (**E**) (Table 2) (3). The results thus strongly point to interaction differences between the bases in the contact pair, with more attractive interactions in the case of **B** and **E** than in the case of **M**. These differences must arise from the presence of fluorine. Three effects may account for this.

C–F···H–N interactions. In the contact pairs of **AB**, **AE**, **CB** and **CE**, the average distance between F at C4 of the base analog and a hydrogen atom of the exocyclic amine group of **A** or **C** is 2.71 ± 0.30 Å. This distance is slightly higher than the sum of the van der Waals radii of F (1.47 Å) and H (1.20 Å) and agrees with distances observed for C–F···H–X (X = N, O) contacts in X-ray structures, where the C–F···H–X contacts have been implied to contribute to stability (10). Covalently bound fluorine hardly ever acts as acceptor for available Bronsted acidic sites due to the low polarizability of fluorine (6–9). However, the strong electronegativity of the fluorine atom and, thus, the strong dipole of the C–F bond (see below) favors dipole–dipole interactions, particularly in appropriately organized systems (10–12,16,17). Apparently, the fluorinated base analog/natural base Watson–Crick-type configurations constitute such well-structured systems so that stabilizing C–F···H–N dipolar interactions occur. Similarly, weak C–F···H–C dipolar interactions were identified recently as stabilizing forces in RNA duplexes that bear fluorine-substituted base analog self-pairs (18).

Figure 4 depicts molecular and bond dipoles of selected bonds of **A** and **M**, **B**, and **E**, respectively, in Watson–Crick-like configurations. The orientation of the *molecular* dipole moments in the **AB** (**AE**) pair is more unfavorable than in the **AM** pair and, hence, does not explain the computed pairing free-energy differences. This agrees with findings from a recent study on interaction energies of hydrogen-bonded **AT** structures determined at the MP2 level (60). At the same time, the study emphasizes that, among various energy components, first-order electrostatic energy, which includes dipole–dipole interactions, plays the most important role in stabilizing base pair configurations (60). Along these lines, differences in the magnitude and orientation of *bond* dipoles of **M** and **B** (**E**) allow for a qualitative explanation of differences in the free-energy curves of **AM** and **AB** (**AE**) at the contact minimum. As such, C–F bonds in **B** (**E**) show dipoles that are at least 1.5 times larger than dipoles of respective C–H bonds in **M**. More importantly, C–F and C–H bond dipoles show reversed orientations, with the orientation of the C–F bond dipole being favorable with respect to the H–N bond dipole of the exocyclic amine group of **A** or **C**.

N···H–C interactions. Due to its electron-withdrawing effect, the presence of fluorine enhances the ability of neighboring functional groups to donate a hydrogen bond (37). The acidity of neighboring hydrogens is also enhanced due to the influence of fluorine (61). Consequently, interactions between N1 (N3) of the natural base and H–C3 (C5) of **B** (**E**) are expected to be stronger than those with **M** due to the presence of two fluorine atoms in *o*- and *o*′-positions, which results in a bond dipole of H–C3 (C5) of **B** (**E**) that is about three times larger than that of H–C3 of **M** (Figure 4). The interactions can be described as weak hydrogen bonds (62), indicated by an N···H distance of on average 2.60 ± 0.20 Å. To further probe this assumption, we calculated a free-energy curve for an **AB*** base pair, where **B*** denotes that the fluorinated base analog is rotated by 180° around the C1-methyl bond compared to the orientation in **AB**. This results in a base pair configuration where N1 of **A** is now opposed by H–C5 instead of H–C3. The pairing free energy for **AB*** was found to be more unfavorable (1.02 kcal mol^{-1}) than for **AB** (0.70 kcal mol^{-1}) (Figure 5). We assume no differences in the C4–F···H–N interactions in both cases, justified by very similar C–F···H–N distances of 2.58 ± 0.28 Å for

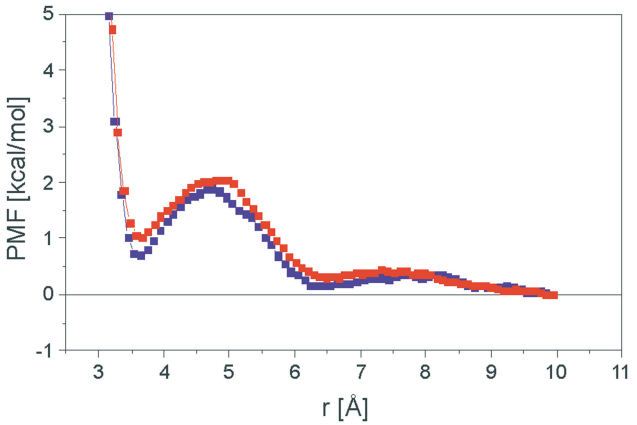


Figure 5. Free-energy profiles as a function of separation distance for the **AB** (blue) and **AB*** (red) configuration.

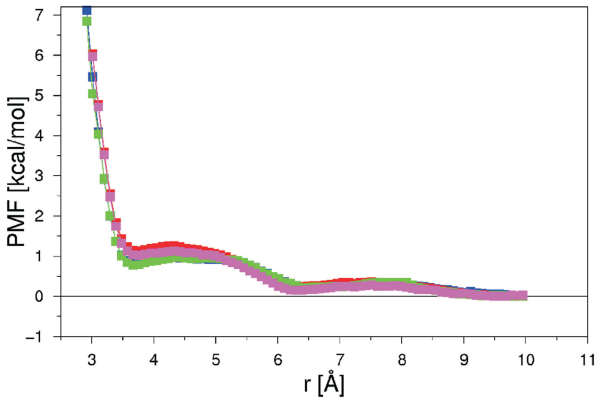


Figure 6. Free-energy profiles as a function of the separation distance for the base pairs **AP** (blue), **CP** (red), **AZ** (green) and **CZ** (magenta).

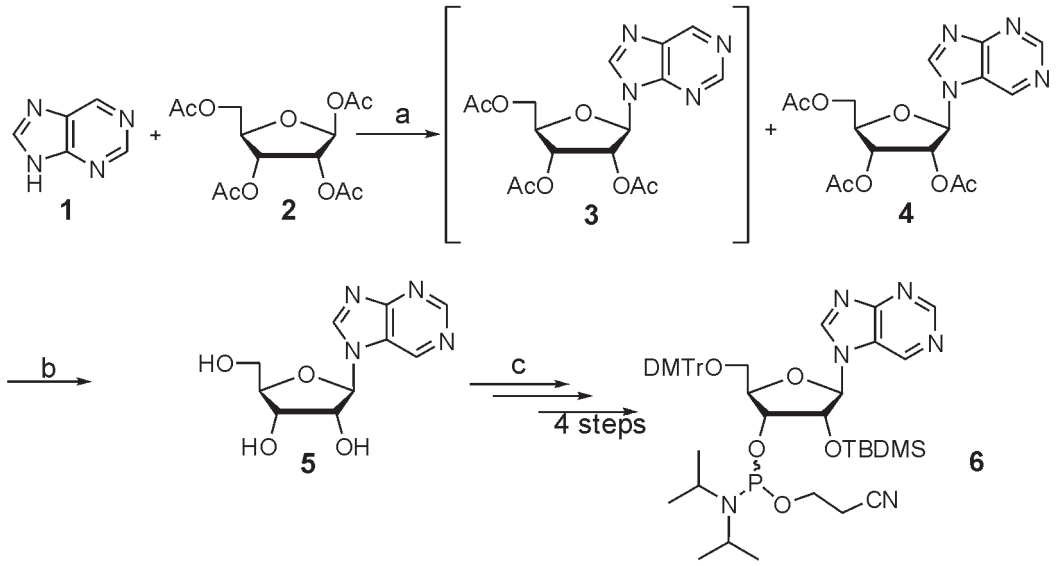
AB and 2.61 ± 0.27 Å for **AB***. The pairing free-energy difference in disfavor of **AB*** thus points to the fact that H–C5 is influenced by two fluorine atoms in *o*- and *p*-position. Compared to an *o*,*o*′-difluoro substitution in the case of **AB**, this results in a reduced electron-withdrawing effect and, hence, a reduced ability to donate a hydrogen-bond. This is also reflected in a reduced bond dipole of H–C5 (1.45 D) compared to H–C3 (1.80 D).

C = O···F–C interactions. In the case of the **GC** Watson–Crick base pair, the carbonyl oxygen of **C** is hydrogen bonded to the exocyclic amino group of **G**, and a loss of this interaction, as in the wobble base pair **GU,** decreases the pairing free energy by 3.1 kcal mol^{-1}. In the case of contact pairs involving **C** and **B** (**E**), the position of the exocyclic amino group of **G** is taken by F attached to C2 (C6) instead. The distances between the carbonyl oxygen and the fluorine atoms are on average 3.70 ± 0.26 Å, considerably larger than the sum of van der Waals radii 2.97 Å of O (1.50 Å) and F (1.47 Å) (63). Intuitively, this arrangement appears unfavorable because two negatively polarized sites face each other. However, the pairing free energies of **AB** (**E**) pairs, where no such arrangement occurs, are almost identical to those of **CB** (**E**) pairs, thus, this does not reflect any unfavorable interactions in the latter cases. This finding is similar to the identification of short F···F contacts in the case of self-pairs of fluorinated bases (18). These contacts did not impair the base pair stability either. Likewise, a strong affinity of a thrombin-inhibitor complex was observed despite a non-covalent O···F interaction between a carbonyl oxygen of a peptide bond and a fluoro-aromatic moiety of the ligand (11,64).

N-heterocyclic universal bases

The above results demonstrate that fluorinated base analogs interact via dipolar interactions with natural bases upon pairing. These interactions originate from the strong electronegativity of the fluorine atom and, thus, the strong dipole of the C–F bond. Nitrogen-containing heterocycles have been proposed as mimics of aromatic rings bearing fluorine atoms, on the account that if C–F is replaced by N both types of compounds display similar electrostatic potentials on their molecular surfaces (37). This is also confirmed by the ParaFrag approach (65), which, in addition to the molecular electrostatic potential, considers the local ionization energy, the local electron affinity, and the local polarizability calculated on isodensity surfaces for similarity comparisons of molecular fragments. According to the ParaFrag similarity score, pyridine is found to be most similar to fluorobenzene (45% similarity) compared to chlorobenzene (27%), pyridine-*N*-oxide (25%), or pyridone (22%).

Following this idea, and using **E** as a template, known to date as the least destabilizing universal base in the context of duplex RNA, we investigated the *N*-heterocyclic analog 7-methyl-7*H*-purine (**P**) with respect to its base pairing properties. To the best of our knowledge, a 7-*N*-linked purine has not yet been applied as a base analog, in contrast to a 9-*N*-linked purine (56,66). The free-energy curves of pairing between **AP** and **CP** are shown in Figure 6, together with curves of the 9-deaza analog of **P**, 5-methyl-5*H*-pyrrolo[3,2-d]pyrimidine (**Z**), paired to **A** and **C**, respectively. **Z** is a more stable analog of **P** (see below). Not surprisingly, the PMFs are very similar for **P** and **Z**, because **P** and **Z** differ only in the structure of the five-membered ring, which is not directly involved in interactions with the natural base. Free energies of pairing are given in Table 1. The difference in the free energy of **P**(**Z**) pairing with either **A** or **C** is −0.29 kcal mol^{-1} (−0.21 kcal mol^{-1}). This difference is similar to the one obtained for **B** (**E**) pairing to **A** or **C**, suggesting that **P** and **Z** will act as universal bases as well. Furthermore, the pairing free energies involving **P** or **Z** are comparable to those obtained for **B** (**E**), but more favorable by at least 0.43 kcal mol^{-1} than the pairing free energy of **M**. This points to the existence of dipolar interactions between **P** (**Z**) and the natural bases and demonstrates that heterocyclic nitrogen indeed can act as a mimic for an aromatic C–F group in these cases. These results led us to suggest using **P** and **Z** as novel universal bases in the context of duplex RNA.



Scheme 3. Synthesis of 1′-deoxy-1′-(7-*H*-purin-7-yl)-β-D-ribofuranose phosphoramidite (**6**). Key: (**a**) BSA = *N,O*–bis(trimethylsilyl)acetamide, TMSOTf = trimethylsilyl trifluoromethanesulfonate, CH_3CN, microwave, 66%; (**b**) NaOMe/MeOH, 97%; (**c**) 13%.

Synthesis

To incorporate **P** and **Z** as base analogs into duplex RNA, phosphoramidite derivatives of the respective ribonucleotides were synthesized as building blocks following a previously developed strategy (3,67).

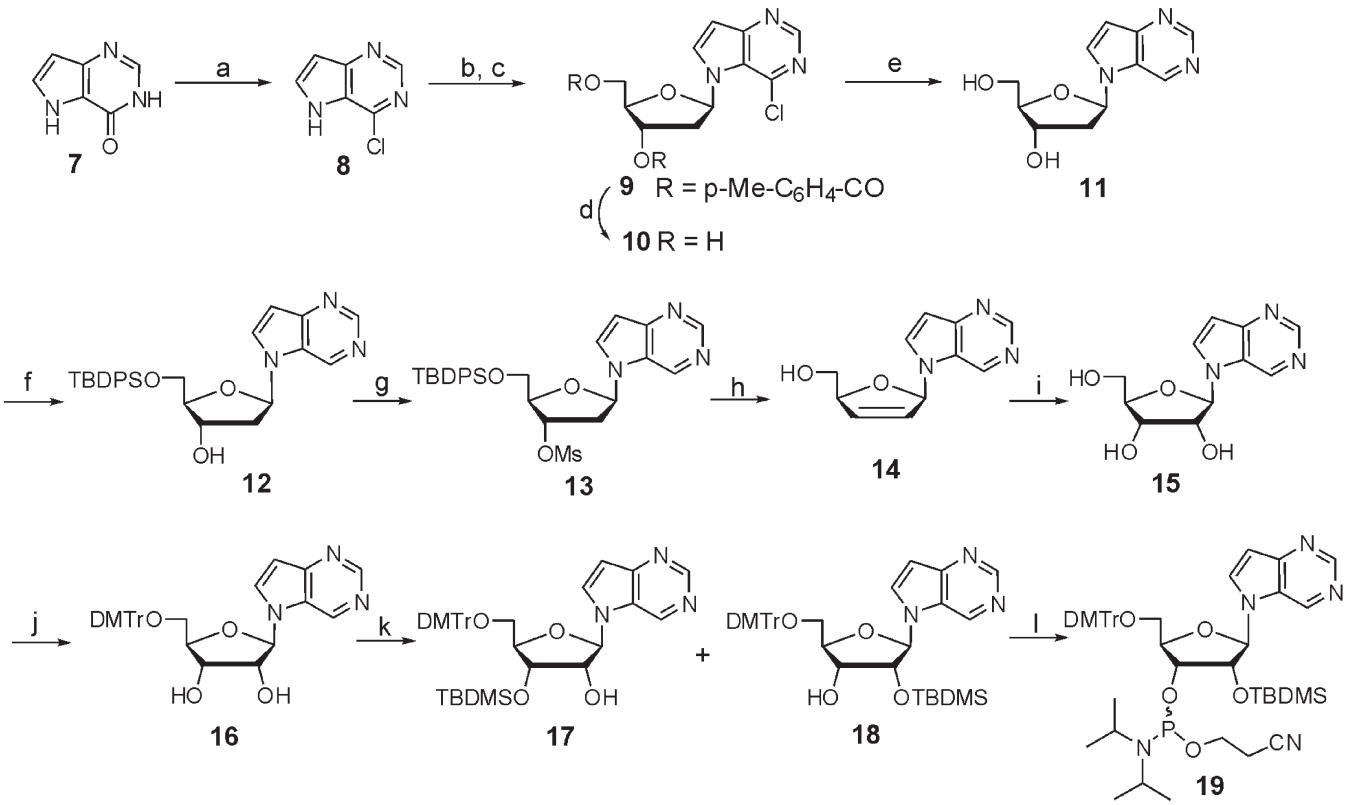
The synthesis of 1′-deoxy-1′-(7*H*-purin-7-yl)-β-D-ribofuranose phosphoramidite **6** (Scheme 3) followed the glycosylation procedure of Vorbrüggen (68). Refluxing of 7*H*-purine (**1**) with *N,O-bis*(trimethylsilyl)acetamide and subsequent reaction of the persilylated base with of 1,2,3,5-tetra-*O*-acetyl-β-D-ribofuranose (**2**) in the presence of the Lewis acid trimethylsilyl trifluoromethanesulfonate afforded a mixture of 2′,3′,5′-tri-*O*-acetyl-1′deoxy-1′-(9*H*-purin-9-yl)- β-D-ribofuranose (**3**) and 2′,3′,5′-tri-*O*-acetyl-1′-deoxy-1′-(7*H*-purin-7-yl)-β-D-ribofuranose (**4**). The deprotection of the acetylated nucleoside (**4**) with a catalytic amount of sodium methoxide furnished 1′-deoxy-1′-(7*H*-purin-7-yl)-β-D-ribofuranose (**5**). For preparation of 1′-deoxy-1′-(7-*H*-purin-7-yl)-β-D-ribofuranose phosphoramidite (**6**) a four-step procedure was followed as previously described (3). See Supplementary Data for further details.

For 1′-deoxy-1′-(9-deaza-7*H*-purin-7-yl)-β-D-ribofuranose (**15**) (Scheme 4) 4,5-dihydro-3*H*-pyrrolo[3,2–d]pyrimidin-4-one (9-deazahypoxanthine) (**7**) was first synthesized as described (69). Chlorination of (**7**) with phosphorus(V)oxychloride yielded 4-chloro-5*H*-pyrrolo[3,2-d]pyrimidine (**8**). For glycosylation, (**8**) was deprotonated with sodium hydride and then reacted with 2′-deoxy-3′,5′-di-*O*-(4-methylbenzoyl)-α-D-ribofuranosyl chloride to give intermediate (**9**). 2′-deoxy-3′,5′-di-*O*-(4-methylbenzoyl)-α-D-ribofuranosyl chloride was synthesized according to Rolland (70). Deprotection of (**9**) with ammonia and catalytic dechlorination of (**10**) afforded the nucleoside (**11**). The selective protection of the 5′-OH group with *tert*-butyldiphenylchlorosilane (TBDPSCl) and subsequent mesylation of the 3′-OH group delivered the completely protected nucleoside (**13**). Tetrabutylammonium fluoride (TBAF) was then used to deprotect the 5′-OH group and eliminate the 3′-OH group in one step. Using osmium tetroxide in catalytic amounts, (**14**) was dihydroxylated in the presence of 4-methylmorpholine-*N*-oxide and water affording 1′-deoxy-1′-(9-deaza-7*H*-purin-7-yl)-β-D-ribofuranose (**15**). The preparation of the final phosphoramidite (**19**) followed a procedure previously described (3). See Supplementary Data for further details.

UV-melting profiles of duplex RNA containing P and Z

With the help of the synthesized phosphoramidites (**6**) and (**19**), nucleotides of **P** and **Z** were incorporated into a 12-mer duplex RNA (5′-CUU UUC XUU CUU-3′ paired with 3′-GAA AAG YAA GAA-5′) at position X. The base analogs were paired with all natural RNA bases at position Y. A system of this type has been previously used by us to investigate thermodynamic properties of duplex RNA by UV-melting measurements (3). In agreement with previous investigations (3,59), the CD spectra of the RNA duplex structures follow the typical curves for an A-form helix (Supplementary Figure S5). Thus, the structure of the 12-mer RNA duplex is not perturbed by incorporation of **P** or **Z,** and all differences determined by UV measurements are a consequence of changes in stacking, solvation, or base pairing interactions and not of structural changes in the RNA duplex. Results of UV-melting experiments are given in Table 2.

Compared to the fluorinated base analogs **B** and **E**, incorporation of **P** and **Z** as universal bases generally results in higher RNA stabilities, as expressed by T_m values between 32.2 and 35.8°C (with the exception of X = **P**, Y = **U**: 30.8°C). As the computed pairing free energies were very similar irrespective of the base analog used (see above), indicating similar interactions with the pairing base, the higher RNA stabilities measured for **P**



Scheme 4. Synthesis of 1′-deoxy-1′-(9-deaza-7*H*-purin-7-yl)-ß-D-ribofuranose phosphoramidite (**19**). Key: (**a**) $POCl_3$, reflux, 3 h, 86%; (**b**) NaH, MeCN, 10 min; (**c**) 2′-deoxy-3′,5′-di-*O*-(4-methylbenzoyl)-α-D-ribofuranosyl chloride, 20 min, 85%; (**d**) NH_3/MeOH, 12 h, 61%; (**e**) H_2, Pd/C, 12 h, 76%; (**f**) TBDPSCl, pyridine, 24 h, 76%; (**g**) MsCl, CH_2Cl_2/pyridine, 12 h, 96%; (**h**) TBAF, THF, 50°C, 2 h, 93%; (**i**) OsO_4, NMO, 20 h, 36%, (**j**) DMTrCl, Et_3N, pyridine, 48 h, 54%, (**k**) TBDMSCl, $AgNO_3$, pyridine/THF, 20 h 32% (**17**), 32% (**18**), (**l**) CEPCl, *sym*-Collidin, 1-methylimidazol, CH_3CN,1 h, 73%.

and **Z** must arise from differences in solvation and stacking interactions. In comparison to **B**, this can be explained in general by stronger stacking interactions observed for larger base scaffolds (71).

Regarding the variation of RNA duplex stabilities either containing **P** or **Z** in response to the paired natural base, both **P** and **Z** exhibit the proposed universal base character, with **Z** showing variations in the T_m values of only 2.7°C. Notably, predicted differences in the free energies for **P** (**Z**) pairing against **A** or **C** of −0.29 (−0.21) kcal mol^{-1} (Table 1) are in close agreement with experimental stability differences of 0.2 (−0.3) kcal mol^{-1} (Table 2).

With respect to a Watson–Crick base pair X = **U**, Y = **A**, for which a T_m value of 37.8°C was measured (Table 2) (3), **P** and **Z** destabilize the 12-mer RNA duplex by at most 5.6 and 4.7°C, respectively. With regard to **Z**, this destabilization is smaller by 4.2°C on average compared with using **E** as a universal base (3). Thus, **Z** is, at present, the least destabilizing universal base in the context of duplex RNA.

CONCLUSION

Molecular dynamics simulations and umbrella sampling were used to calculate pairing free energies between natural nucleobases and fluorinated base analogs in a Watson–Crick-like orientation in aqueous solution. Compared to pairing free energies of Watson–Crick base pairs **AU** (−3.07 kcal mol^{-1}) and **GC** (−5.98 kcal mol^{-1}), base pairing involving fluorinated base analogs is unfavorable as demonstrated by pairing free energies of 0.55–1.07 kcal mol^{-1}. These results compare favorably with duplex RNA stability changes observed experimentally upon replacing the natural base **U** in the Watson–Crick pair **AU** by a fluorinated analog. The fluorinated base analogs exhibit universal base pairing properties as shown by similar free energies of pairing to different natural nucleobases, in agreement with experimental studies (3). Compared to Watson–Crick base pairs, the location of the contact minima between natural nucleobases and fluorinated base analogs is shifted by ∼0.7 Å to larger distances. As in the case of the Watson–Crick base pairs, the PMFs of the natural nucleobase/fluorinated base analog pairs show a secondary minimum, where two water molecules bridge the gap between the bases.

A decomposition of the pairing free energies into enthalpic and entropic components revealed similar contributions for both a Watson–Crick complex (**AU**) and one involving a fluorinated base analog (**AB**) when it comes to expelling the bridging waters. Upon formation

of the contact pairs, however, enthalpic and entropic components showed fundamental differences. The Watson–Crick base pair formation is enthalpically favored but entropically disfavored, as expected by the formation of hydrogen bonds, which restricts the configuration space of the bases. In contrast, formation of a natural nucleobase/fluorinated base analog pair has an unfavorable enthalpy component, which is partially compensated by a favorable entropy contribution, demonstrating that both bases are less strongly locked in. Overall, these findings indicate a pronounced enthalpy–entropy compensation during the process of base pair formation.

Pairing free energies calculated for a non-fluorinated base analog/natural nucleobase pair (**AM**), in comparison to a fluorinated base analog/natural nucleobase pair (**AB**), strongly point to interaction differences between the bases in the contact pairs, with more attractive interactions in the case of **B** than in the case of **M**. First, these differences were attributed to stabilizing C–F···H–N dipolar interactions, similar to weak C–F···H–C interactions identified recently as stabilizing forces in RNA duplexes containing fluorinated base analog self-pairs (18). Second, due to the presence of fluorine, the ability of neighboring functional groups to donate a hydrogen bond is enhanced, resulting in stronger N···H–C hydrogen bonds between the bases, in particular, in the case of the *o*,*o*′-difluoro substituted base analogs **B** and **E**. Third, C = O···F–C contacts were not found to be unfavorable in the case of **CB** (**E**) pairs.

Based on these findings, 7-methyl-7*H*-purine (**P**) and its 9-deaza analog (**Z**) were suggested as members of a new class of non-fluorinated base analogs. Here, nitrogen-containing heterocycles mimic aromatic rings bearing fluorine atoms. Computed pairing free energies of **AP** or **AZ** pairs are similar to those of base pairs containing fluorinated base analogs (**AB**, **AE**), indicating the existence of dipolar interactions between **P** (**Z**) and natural nucleobases. To confirm these results, phosphoramidite derivatives of the respective ribonucleotides were synthesized as building blocks and incorporated in a 12-mer duplex RNA. Duplex RNA stabilities determined by UV-melting experiments revealed **Z** as the least destabilizing universal base in the context of RNA known to date. To the best of our knowledge, this is the first experimental validation of nitrogen-containing heterocycles as bioisosteres of aromatic rings bearing fluorine atoms.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

The authors thank D. Schade (University of Kiel) for critically reading the Synthesis part, D. Cahsman for crtically reading the manuscript, and Harald Mauser (Hoffmann-La Roche, Basel) for providing the ParaFrag results. They acknowledge support by the Frankfurt Center for Scientific Computing (CSC).

FUNDING

Deutsche Forschungsgemeinschaft via the SFB project 579, 'RNA–ligand interactions'. Funding for open access charges: Institutional funding from Goethe-University and Heinrich-Heine-University.

Conflict of interest statement. None declared.

REFERENCES

1. Schweitzer,B.A. and Kool,E.T. (1994) Aromatic nonpolar nucleosides as hydrophobic isosteres of pyrimidines and purine nucleosides. *J. Org. Chem.*, **59**, 7238–7242.
2. Loakes,D. (2001) Survey and summary: The applications of universal DNA base analogues. *Nucleic Acids Res.*, **29**, 2437–2447.
3. Parsch,J. and Engels,J.W. (2002) C-F...H-C hydrogen bonds in ribonucleic acids. *J. Am. Chem. Soc.*, **124**, 5664–5672.
4. Lai,J.S. and Kool,E.T. (2004) Selective pairing of polyfluorinated DNA bases. *J. Am. Chem. Soc.*, **126**, 3040–3041.
5. Somoza,A., Chelliserrykattil,J. and Kool,E.T. (2006) The roles of hydrogen bonding and sterics in RNA interference. *Angew. Chem. Int. Ed.*, **45**, 4994–4997.
6. Murray-Rust,P., Stallings,W.C., Monti,C.T., Preston,R.K. and Glusker,J.P. (1983) Intermolecular interactions of the carbon-fluorine bond: the crystallographic environment of fluorinated carboxylic acids and related structures. *J. Am. Chem. Soc.*, **105**, 3206–3214.
7. Dunitz,J. and Taylor,R. (1997) Organic fluorine hardly ever accepts hydrogen bonds. *Chem. Eur. J.*, **3**, 89–98.
8. Dunitz,J. (2004) Organic fluorine: odd man out. *ChemBioChem*, **5**, 614–621.
9. Kool,E. and Sintim,H. (2006) The difluorotoluene debate-a decade later. *Chem. Commun.*, 3665–3675.
10. O'Hagan,D. (2008) Understanding organofluorine chemistry. An introduction to the C-F bond. *Chem. Soc. Rev.*, **37**, 308–319.
11. Olsen,J.A., Banner,D.W., Seiler,P., Obst-Sander,U., D'Arcy,A., Stihle,M., Müller,K. and Diederich,F. (2003) A fluorine scan of thrombin inhibitors to map the fluorophilicity/fluorophobicity of an enzyme active site: evidence for C-F...C = O interactions. *Angew. Chem. Int. Ed.*, **42**, 2507–2511.
12. Olsen,J.A., Banner,D.W., Seiler,P., Wagner,B., Tschopp,T., Obst-Sander,U., Kansy,M., Müller,K. and Diederich,F. (2004) Fluorine interactions at the thrombin active site: protein backbone fragments H-C(alpha)-C = O comprise a favorable C-F environment and interactions of C-F with electrophiles. *ChemBioChem*, **5**, 666–675.
13. Olsen,J., Seiler,P., Wagner,B., Fischer,H., Tschopp,T., Obst-Sander,U., Banner,D.W., Kansy,M., Müller,K. and Diederich,F. (2004) A fluorine scan of the phenylamidinium needle of tricyclic thrombin inhibitors: effects of fluorine substitution on pKa and binding affinity and evidence for intermolecular C-F...CN interactions. *Org. Biomol. Chem.*, **2**, 1339–1352.
14. Müller,K., Faeh,C. and Diederich,F. (2007) Fluorine in pharmaceuticals: looking beyond intuition. *Science*, **317**, 1881–1886.
15. Thalladi,V.R., Weiss,H.C., Blaser,D., Boese,R., Nangia,A. and Desiraju,G.R. (1998) C-H...F interactions in the crystal structures of some fluorobenzenes. *J. Am. Chem. Soc.*, **120**, 8702–8710.
16. Desiraju,G.R. (2002) Hydrogen bridges in crystal engineering: interactions without borders. *Acc. Chem. Res.*, **35**, 565–573.
17. Reichenbächer,K., Süss,H. and Hulliger,J. (2005) Fluorine in crystal engineering – "the little atom that could". *Chem. Soc. Rev.*, **34**, 22–30.
18. Kopitz,H., Živcović,A., Engels,J.W. and Gohlke,H. (2008) Determinants of the unexpected stability of RNA fluorobenzene self pairs. *ChemBioChem*, **9**, 2619–2622.
19. Turner,D. (1996) Thermodynamics of base pairing. *Curr. Opin. Struct. Biol.*, **6**, 299–304.

20. Stofer,E., Chipot,C. and Lavery,R. (1999) Free energy calculations of Watson-Crick base pairing in aqueous solution. *J. Am. Chem. Soc.*, **121**, 9503–9508.
21. Freier,S.M., Sugimoto,N., Sinclair,A., Alkema,D., Neilson,T., Kierzek,R., Caruthers,M.H. and Turner,D.H. (1986) Stability of XGCGCp, GCGCYp, and XGCGCYp helixes: an empirical estimate of the energetics of hydrogen bonds in nucleic acids. *Biochemistry*, **25**, 3214–3219.
22. Turner,D., Sugimoto,N., Kierzek,R. and Dreiker,S. (1987) Free energy increments for hydrogen bonds in nucleic acid base pairs. *J. Am. Chem. Soc.*, **109**, 3783–3785.
23. Tanner,N.K. and Cech,T.R. (1987) Guanosine binding required for cyclization of the self-splicing intervening sequence ribonucleic acid from *Tetrahymena thermophila*. *Biochemistry*, **26**, 3330–3340.
24. Kierzek,E., Pasternak,A., Pasternak,K., Gdaniec,Z., Yildirim,I., Turner,D. and Kierzek,R. (2009) Contributions of stacking, preorganization, and hydrogen bonding to the thermodynamic stability of duplexes between RNA and 2'-*O*-methyl RNA with locked nucleic acids. *Biochemistry*, **48**, 4377–4387.
25. Vesnaver,G. and Breslauer,K.J. (1991) The contribution of DNA single-stranded order to the thermodynamics of duplex formation. *Proc. Natl Acad. Sci. USA*, **88**, 3569–3573.
26. Levitt,M. (1978) How many base-pairs per turn does DNA have in solution and in chromatin? Some theoretical calculations. *Proc. Natl Acad. Sci. USA*, **75**, 640–644.
27. Arora,N. and Jayaram,B. (1998) Energetics of base pairs in B-DNA in solution: an appraisal of potential functions and dielectric treatments. *J. Phys. Chem. B*, **102**, 6139–6144.
28. Cubero,E., Sherer,E., Luque,J., Orozco,M. and Laughton,C. (1999) Observation of spontaneous base pair breathing events in the molecular dynamics simulation of a difluorotoluene-containing DNA oligonucleotide. *J. Am. Chem. Soc.*, **121**, 8653–8654.
29. Meyer,M. and Sühnel,J. (1997) Quantum-chemical ab initio study on the adenine-difluorotoluene complex–a mimic for the adenine-thymine base pair. *J. Biomol. Struct. Dyn.*, **15**, 619–624.
30. Ryjácek,F., Kratochvíl,M. and Hobza,P. (1999) Adenine – 2,4-difluorotoluene (modified base) pair: potential and free-energy surfaces: non-empirical and empirical potential studies. *Chem. Phys. Lett.*, **313**, 393–398.
31. Cubero,E., Laughton,C., Luque,J. and Orozco,M. (2000) Molecular dynamics study of oligonucleotides containing difluorotoluene. *J. Am. Chem. Soc.*, **122**, 6891–6899.
32. Florian,J., Goodman,M. and Warshel,A. (2000) Free-energy perturbation calculations of DNA destabilization by base substitutions: the effect of neutral guanine.thymine, adenine.cytosine and adenine.difluorotoluene mismatches. *J. Phys. Chem. B*, **104**, 10092–10099.
33. Bats,J.W., Parsch,J. and Engels,J.W. (2000) 1-deoxy-1-(4-fluorophenyl)-beta-D-ribofuranose, its hemihydrate, and 1-deoxy-1-(2,4-difluorophenyl)-beta-D-ribofuranose: structural evidence for intermolecular C-H…F-C interactions. *Acta Crystallogr., Sect. C: Cryst. Struct. Commun.*, **56(Pt 2)**, 201–205.
34. Živković,A. and Engels,J.W. (2003) Synthesis of modified RNA-oligonucleotides for structural investigations. *Nucleosides Nucleotides Nucl. Acids*, **22**, 1167–1170.
35. Klöpffer,A.E. and Engels,J.W. (2004) Synthesis of 2′-aminoalkyl-substituted fluorinated nucleobases and their influence on the kinetic properties of hammerhead ribozymes. *ChemBioChem*, **5**, 707–716.
36. Živković,A. and Engels,J.W. (2007) Florobenzene as artificial nucleobases-base pairing and stacking interactions. *Nucleosides Nucleotides Nucl. Acids*, **26**, 559–562.
37. Bégué,J.-P. and Bonnet-Delpon,D. (2008) *Bioorganic and Medicinical Chemistry of Fluorine*, 1st edn. John Wiley & Sons, Inc., Hoboken, New Jersey, USA.
38. Bayly,C.I., Cieplak,P., Cornell,W.D. and Kollman,P.A. (1993) A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: the RESP model. *J. Phys. Chem.*, **97**, 10269–10280.
39. Cornell,W.D., Cieplak,P., Bayly,C., Gould,I.R., Merz,K.M., Ferguson,D.M., Spellmeyer,D.C., Fox,T., Caldwell,J.W. and Kollman,P.A. (1995) A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *J. Am. Chem. Soc.*, **117**, 5179–5197.
40. Müller,U., Schübel,H., Sprinzl,M. and Heinemann,U. (1999) Crystal structure of acceptor stem of tRNA(Ala) from Escherichia coli shows unique G.U wobble base pair at 1.16 A resolution. *RNA*, **5**, 670–677.
41. Torrie,G.M. and Valleau,J.P. (1977) Nonphysical sampling distributions in Monte Carlo free-energy estimation - Umbrella sampling. *J. Comput. Phys.*, **23**, 187–199.
42. Case,D.A., Cheatham,T., Darden,T., Gohlke,H., Luo,R., Merz,K.M., Onufriev,A., Simmerling,C., Wang,B. and Woods,R. (2005) The Amber biomolecular simulation programs. *J. Comput. Chem.*, **26**, 1668–1688.
43. Šponer,J., Jurečka,P. and Hobza,P. (2004) Accurate interaction energies of hydrogen-bonded nucleic acid base pairs. *J. Am. Chem. Soc.*, **126**, 10142–10151.
44. Šponer,J.E., Špačková,N., Kulhánek,P., Leszczynski,J. and Šponer,J. (2005) Non-Watson-Crick base pairing in RNA. quantum chemical analysis of the cis Watson-Crick/sugar edge base pair family. *J. Phys. Chem. A*, **109**, 2292–2301.
45. Šponer,J.E., Špačková,N., Leszczynski,J. and Šponer,J. (2005) Principles of RNA base pairing: structures and energies of the trans Watson-Crick/sugar edge base pairs. *J. Phys. Chem. B*, **109**, 11399–11410.
46. Jorgensen,W. (1982) Revised TIPS for simulations of liquid water and aqueous solutions. *J. Chem. Phys.*, **77**, 4156–4163.
47. Berendsen,H.J.C., Postma,J.P.M., van Gunsteren,W.F., Dinola,A. and Haak,J.R. (1984) Molecular dynamics with coupling to an external bath. *J. Chem. Phys.*, **81**, 3684–3690.
48. Ryckaert,J.P., Ciccotti,G. and Berendsen,H.J.C. (1977) Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J. Comput. Phys.*, **23**, 327.
49. Darden,T., York,D. and Pedersen,L. (1993) Particle mesh Ewald: an N*log(N) method for Ewald sums in large systems. *J. Chem. Phys.*, **98**, 10089–10092.
50. Bouzida,D., Kumar,S. and Swendsen,R.H. (1992) Efficient Monte Carlo methods for the computer simulation of biological molecules. *Phys. Rev. A*, **45**, 8894–8901.
51. Efron,B. and Tibshirani,R. (1994) *An Introduction to the Bootstrap*. Chapman & Hall/CRC, Boca Raton, Florida, USA.
52. MacCallum,J.L., Mukhopadhyay,P., Luo,H. and Tieleman,D.P. (2003) In Sénéchal,D. (ed.), *Proceedings of the 17th Annual International Symposium on High Performance Computing Systems and Applications and the OSCAR Symposium*. NRC Research Press, Ottawa, Canada, Sherbrooke, Canada, pp. 115–122.
53. Varani,G. and McClain,W.H. (2000) The G×U wobble base pair. A fundamental building block of RNA structure crucial to RNA function in diverse biological systems. *EMBO Rep.*, **1**, 18–23.
54. Buckin,V.A. (1988) Hydration of nucleic bases in dilute aqueous solutions. Apparent molar adiabatic and isothermal compressibilities, apparent molar volumes and their temperature slopes at 25 degrees C. *Biophys. Chem.*, **29**, 283–292.
55. Gao,J. (1994) The hydration and solvent polarization effects of nucleotide bases. *Biophys. Chem.*, **51**, 253–261.
56. Kool,E., Morales,J. and Guckian,K. (2000) Mimicking the structure and function of DNA: Insights into DNA stability and replication. *Angew. Chem. Int. Ed.*, **39**, 990–1009.
57. Šponer,J., Leszczynski,J. and Hobza,P. (1996) Hydrogen bonding and stacking of DNA bases: a review of quantum-chemical ab initio studies. *J. Biomol. Struct. Dyn.*, **14**, 117–135.
58. Barsky,D., Kool,E.T. and Colvin,M.E. (1999) Interaction and solvation energies of nonpolar DNA base analogues and their role in polymerase insertion fidelity. *J. Biomol. Struct. Dyn.*, **16**, 1119–1134.
59. Zacharias,M. and Engels,J.W. (2004) Influence of a fluorobenzene nucleobase analogue on the conformational flexibility of RNA studied by molecular dynamics simulations. *Nucleic Acids Res.*, **32**, 6304–6311.
60. Ran,J. and Hobza,P. (2009) Nature of bonding in nine planar hydrogen-bonded adenine···thymine base pairs. *J. Phys. Chem. B*, **113**, 2933–2936.
61. Smart,B. (2001) Fluorine substituent effects (on bioactivity). *J. Fluorine Chem.*, **109**, 3–11.

62. Jeffrey,G.A. (1997) *An Introduction to Hydrogen Bonding*. Oxford University Press, New York.
63. Bondi,A. (1964) Van der Waals volumes and radii. *J. Phys. Chem.*, **68**, 441–451.
64. Hof,F., Scofield,D.M., Schweizer,W.B. and Diederich,F. (2004) A weak attractive interaction between organic fluorine and an amide group. *Angew. Chem. Int. Ed.*, **43**, 5056–5059.
65. Jakobi,A.-J., Mauser,H. and Clark,T. (2008) ParaFrag—an approach for surface-based similarity comparison of molecular fragments. *J. Mol. Modeling*, **14**, 547–558.
66. Eritja,R., Horowitz,D., Walker,P., Ziehler,M., Boosalis,M., Goodman,M., Itakura,K. and Kaplan,B. (1986) Synthesis and properties of oligonucleotides containing 2'-deoxynebularine and 2'-deoxyxanthosine. *Nucleic Acids Res.*, **14**, 8135–8153.
67. Parsch,J. and Engels,J.W. (2000) Synthesis of fluorobenzene and benzimidazole nucleic-acid analogues and their influence on stability of RNA duplexes. *Helv. Chim. Acta*, **83**, 1791–1808.
68. Vorbrüggen,H. and Höfle,G. (1981) On the mechanism of nucleoside synthesis. *G. Chem. Ber.*, **114**, 1256–1268.
69. Furneaux,R.H. and Tyler,P.C. (1999) Improved synthesis of 3*H*,5*H*-pyrrolo[3,2-d]pyrimidines. *J. Org. Chem.*, **64**, 8411–8412.
70. Rolland,V., Kotera,M. and Lhomme,J. (1997) Convenient preparation of 2-deoxy-3,5-di-O-*p*-toluolyl-α-D-*erythro*-pentofuranosyl chloride. *Synthetic Comm.*, **27**, 3505–3511.
71. Guckian,K., Schweitzer,B., Ren,R., Sheils,C., Paris,P., Tahmassebi,D. and Kool,E. (1996) Experimental measurement of aromatic stacking affinities in the context of duplex DNA. *J. Am. Chem. Soc.*, **118**, 8182–8183.