Supporting Information

Determinants of the unexpected stability of RNA fluorobenzene self pairs

Hannes Kopitz, Aleksandra Živković, Joachim W. Engels*, Holger Gohlke*

1 Experimental methods

1.1 Synthesis of the benzene and the fluorinated benzene nucleosides

The anhydrous solvents were obtained from *Fluka* and used without further purification. Dry MeCN ($H_2O < 30$ ppm) for the phosphitylation reaction was purchased from *PerSeptive Biosystems*. Flash column chromatography (FC): silica gel 60 (40 – 63 µm) from *Merck*. TLC: silica gel 60 F_{254} plates from *Merck*; HPLC: anion-exchange column NucleoPac PA-100 from *Dionex*; desalting Sephadex-G25 columns from *Pharmacia*. UV/melting profiles: UV/vis spectrophotometer Cary-1 from *Varian*, Cary temperature controller, 10 mm cuvette. CD spectra: Spectropolarimeter 710 from *Jasco*. NMR: Spectrometers AM250 (1H, 13C) and AMX400 (1H, 13C, 31P) from *Bruker*; γ in ppm, J in Hz. MS: MALDI-TOF spectrometer Voyager DE from *PerSeptive Biosystems*; ESI: electron spray ionisation.

1.1.1 Synthesis of the benzene, 4-fluorobenzene, 2,4-difluorobenzene and 2,4,5-trifluorobenzene nucleosides

The syntheses of the benzene, 4-fluorobenzene, and 2,4-difluorobenzene nucleosides have been described in ref.[1,2], the synthesis of the 2,4,5-trifluorobenzene nucleoside has been described in ref.[3].

1.1.2 Synthesis of the 2,3,4,5-Tetrafluorobenzene nucleoside

2',3',5'-Tri-*O*-benzyl-1'-deoxy-1'-(2,3,4,5-tetrafluorophenyl)-β-D-ribofuranose To the solution that contains 1 g (5 mmol) 1-bromo-2,3,4,5-tetrafluorobenzene in 10 ml abs. diethyl ether at –78 °C under argon over 10 min 3.2 ml of 1.6 M

solution of n-butyl lithium in *n*-hexane were added. After 20 min at –78 °C a solution of 2,3,5-tri-*O*-benzyl-ribono-γ-lacton (2.0 g, 4.8 mmol) in 30 ml abs. diethyl ether was added over 30 min and stirred at –78 °C for 1 h. Afterwards the reaction mixture was allowed to warm up to -20 °C over 2 h. The reaction was stopped by adding 5 ml of water and three times extracted with ether. The combined organic phases were dried over $MgSO_4$ and evaporated to dryness. Residual oil was immediately dissolved in 30 ml of methylene chloride, cooled to -78 °C, and 1.3 ml (10.0 mmol) bortrifluoridethyletherat and 1.6 ml (10.0 mmol) triethylsilane were added. The reaction mixture was left at -78 °C and allowed to warm up to 10 °C over night. The reaction was stopped by adding 10 ml of saturated water solution of $NaHCO_3$ and extracted three times with methylene chloride. The combined organic phases were dried over $MgSO_4$ and evaporated under reduced pressure. Further purification was done by FC using *n*-hexane/ethyl acetate 4:1 as eluent. The product was obtained as light orange solid. Yield: 1.6 g (67 %); TLC: (*n*-hexane/ethyl acetate 4:1) R_f = 0.35; ^{1}H-NMR: γ [ppm] (400 MHz, DMSO-d_6): 7.32 (m, 16H, H_{ar}, 5H); 5.19 (d, J = 4.5 Hz, 1H, 1'H); 4.58 (m, 6H, CH_2-benzyl); 4.23 (q, J = 3.9 Hz, 1H, 4'H); 4.15 (m, 2H, 2'H, 3'H); 3.70 (m, 2H, 5'H); ^{13}C-NMR: γ [ppm] (67.9 MHz, DMSO-d_6): 162.51 (C4); 158.87 (C2); 138.18 (C_{ar}); 138.10 (C_{ar}); 137.96 (C_{ar}); 129.36 (C6); 128.22 (C_{ar}); 128.09 (C_{ar}); 127.70 (C_{ar}); 127.50 (C_{ar}); 127.42 (C_{ar}); 127.11 (C_{ar}); 123.61 (C1); 111.30 (C5); 103.62 (C3); 82.10 (C1'); 80.72 (C4'); 76.84 (C2'); 75.96 (C3'); 72.40 (CH_2-benzyl); 71.04 (CH_2-benzyl); 69.73 (C5'); ESI(+): m/z 573.86 ($[M+Na]^+$);

1'-Deoxy-1'-(2,3,4,5-tetrafluorophenyl)-β-D-ribofuranose 1.1 g (2.0 mmol) 2',3',5'-Tri-*O*-benzyl-1'-deoxy-1'-(2,3,4,5-tetrafluorophenyl)-β-D-ribofuranose was dissolved in 40 ml abs. ethanol, and 20 ml cyclohexene and 400 mg palladiumhydroxide (20 %) on carbon was added. The reaction mixture was refluxed for 5 h. Then the reaction mixture was filtrated over celite and the filtrate was evaporated under reduced pressure. The purification was done by FC with methylene chloride/methanol 95:5, as eluent. The product was obtained as colorless solid. For analytical purposes the product was crystallized from water and methanol. Yield: 0.50 g (89.3 %); TLC: (CH_2Cl_2/MeOH 95:5) R_f = 0.33 ; ^{1}H-NMR: γ [ppm] (270 MHz, DMSO-d_6): 7.62 (m, 1H, 5H); 5.17 (d, J = 6.05 Hz, 1H, 1'-OH); 4.95 (m, 3H, 2'-OH, 3'-OH, 5'-OH); 3.98 (m, 3H, 2'H, 3'-H, 4'H); 3.55 (m, 2H, 5'H); ^{13}C-NMR: γ [ppm] (67.9 MHz, DMSO-d_6): 162.58 (C4); 158.94 (C2); 129.67 (C6); 124.40 (C1); 114.74(C5); 103.42 (C3); 84.59 (C1'); 77.08 (C4'); 76.65 (C2'); 70.94 (C3'); 61.59 (C5'); ^{19}F-NMR: γ [ppm] (254.2 MHz, DMSO-d_6): –139.75 (1F, 2F); –143.93 (1F, 3F); 157.30 (1F, 1F); –158.14 (1F, 4F); ESI(–): m/z 280.9 ($[M-H]^-$);

5'-*O*-(4,4'-Dimethoxytriphenylmethyl)-1'-deoxy-1'-(2,3,4,5-tetrafluorophenyl)-β-D-ribofuranose 0.5 g (1.78 mmol) 1'-Deoxy-1'-(2,3,4,5-tetrafluorophenyl)-β-D-ribofuranose was dissolved in 20 ml abs. pyridine, and 0.84 ml

(6.0 mmol) triethylamine and 0.80 g (2.30 mmol) 4,4'-dimethoxytriphenylmethylchloride were added. The reaction mixture was stirred under argon at RT over night. The reaction was stopped by adding of 3 ml of methanol and saturated water solution of $NaHCO_3$. It was extracted with methylene chloride three times, organic phases were collected and dried over $MgSO_4$ and then evaporated to dryness. The product was twice co-evaporated with toluene. Further purification was done by FC with methylene chloride/methanol 98:2 as eluent. The product was obtained as yellow foam. Yield: 0.86 g (82.7 %); TLC: (CH_2Cl_2/MeOH 95:5) Rf = 0.48; ^{1}H-NMR: γ [ppm] (250 MHz, DMSO-d_6): 7.43 – 6.82 (m, 14H, H_{ar}, 5H); 5.32 (d, J = 5.36 Hz, 1H, 2'-OH); 5.05 (d, J = 5.2 Hz, 1H, 3'-OH); 4.70 (d, J = 5.40 Hz, 1H, 1'H); 3.97 (m, 3H, 2'H, 3'H, 4'H); 3.804 (s, 6H, OCH_3); 3.20 (m, 2H, 5'H); ^{19}F-NMR: γ [ppm] (254.2 MHz, DMSO-d_6): –139.70 (1F, 2F); –143.28 (1F, 3F); 156.81 (1F, 1F); –157.87 (1F, 4F); ESI(+): m/z 585.2 ($[M+H]^+$);

5'-*O*-(4,4'-Dimethoxytriphenylmethyl)-2'-*O-tert.*-butyldimethylsilyl-1'-deoxy-1'-(2,3,4,5-tetrafluorophenyl)-β-D-ribofuranose 800 mg (1.37 mmol) 5'-*O*-(4,4'-Dimethoxytriphenylmethyl)-1'-deoxy-1'-β-D-(2,3,4,5-tetrafluorophenyl)-ribofuranose were dissolved in 20 ml of 1:1 mixture of THF/pyridine, and 260 mg (1.50 mmol) $AgNO_3$ and 2.0 ml (2.0 mmol) 1 M *tert.*-butyldimethylsilylchloride solution in THF were added. The reaction mixture was stirred for 20 h at RT under argon. Adding 10 ml of saturated water $NaHCO_3$-solution stopped the reaction. Precipitated AgCl was filtered over celite, and the filtrate was extracted with methylene chloride three times. The collected organic phases were dried over $MgSO_4$ and evaporated to dryness. The crude product was co-evaporated with toluene twice. Further purification of the product was done by HPLC (MN Nucleoprep 100-20 of *Macherey-Nagel*, *n*-hexane/isopropyl acetate 3:2). The product (*slow-Isomer*) was obtained as a white foam. Yield: 203 mg (21 %); TLC: (CH_2Cl_2) R_f = 0.18; ^{1}H-NMR: γ [ppm] (270 MHz, DMSO-d_6): 7.46 – 6.89 (m, 14H, H_{ar}, 5H); 5.03 (d, J = 4.0 Hz, 1H, 3'-OH); 4.84 (d, J = 5.7 Hz, 1H 1'H); 4.00 (m, 2H, 3'H, 4'H); 3.93 (m, 1H, 2'H); 3.74 (s, 6H, OCH_3); 3.21 (m, 2H, 5'H); 0.80 (s, 9H, $SiC(CH_3)_3$); 0.02 ($SiCH_3$); 0.01 ($SiCH_3$); ^{19}F-NMR: γ [ppm] (254.2 MHz, DMSO-d_6): –139.45 (1F, 2F); –143.23 (1F, 3F); 156.87 (1F, 1F); –157.55 (1F, 4F); ESI(–): m/z 697.0 ($[M-H]^-$);

5'-*O*-(4,4'-Dimethoxytriphenylmethyl)-3'-*tert.*-butyldimethylsilyl-1'-deoxy-1'-(2,3,4,5-tetrafluorophenyl)-β-D-ribofuranose 5'-*O*-(4,4'-Dimethoxytriphenylmethyl)-3'-*O-tert.*-butyldimethylsilyl-1'-deoxy-1'-(2,3,4,5-tetrafluorophenyl)-β-D-ribofuranose was obtained as a side product and faster migrating isomer while synthesizing 5'-*O*-(4,4'-Dimethoxytriphenylmethyl)-2'-*O-tert.*-butyldimethylsilyl-1'-deoxy-1'-(2,4-difluorophenyl)-β-D-ribofuranose. Yield: 344 mg (36.0 %); TLC: (CH_2Cl_2) R_f = 0.19; ^{1}H-NMR: γ [ppm] (270 MHz, DMSO-d_6): 7.47 – 6.88 (m, 14H, H_{ar}, 5H); 5.19 (d, J = 4.0 Hz, 1H, 2'OH); 4.98 (d, J = 4.0 Hz, 1H, 1'H); 4.02 (m, 1H, 3'H); 3.94 (m, 1H; 2'H); 3.74 (s, 6H, OCH_3); 3.29 (m, 1H, 4'H); 3.14

(m, 2H, 5'H); 0.76 ($SiC(CH_3)_3$); 0.02 ($SiCH_3$); –0.08 ($SiCH_3$); ^{19}F-NMR: γ [ppm] (254.2 MHz, DMSO-d_6): –139.71 (1F, 2F); –143.17 (1F, 3F); 156.75 (1F, 1F); –157.76 (1F, 4F); ESI(–): m/z 697.0 ($[M-H]^-$);

3'-*O*-(2-Cyanethoxydiisopropylphosphin)-1'-deoxy-5'-*O*-(4,4'-dimethoxytriphenylmethyl)-1'-(2,3,4,5-tetrafluorophenyl)-2'-*O*-*tert.*-butyldimethylsilyl-β-D-ribofuranose 200 mg (0.29 mmol) 5'-*O*-(4,4'-Dimethoxytriphenylmethyl)-2'-*O*-*tert.*-butyldimethylsilyl-1'-deoxy-1'-('-(2,3,4,5-tetrafluorophenyl))-β-D-ribofuranose were dissolved in 10 ml abs. acetonitrile, and 380 µl (3.0 mmol) sym. collidine and 10 µl (0.13 mmol) 1-methylimidazole were added. The reaction mixture was cooled in an icebath to 0 °C and 110 µl (0.36 mmol) 2-cyanethyldiisopropylchlorophosphoramidite were added. The reaction was stirred for 15 min at 0 °C and for 1 h at RT. The reaction was stopped by adding 10 ml of 0.01 M citric acid and three times extracted with methylene chloride. Combined organic phases were washed twice with 0.01 M citric acid, dried over $MgSO_4$, and evaporated to dryness. Purification was done by FC with *n*-hexane/ethyl acetate, 4:1, as eluent. The product (mixture of two diastereomers) was obtained as white foam. Yield: 200 mg (77.0 %); TLC: (*n*-hexane/ethylacetat 4:1) R_f = 0.34; 0.36; ^{1}H-NMR: γ [ppm] (270 MHz, DMSO-d_6): 7.67 (m, 2H, 6H); 7.53 – 6.77 (m, 30H, H_{ar}, 3H, 5H); 5.14, 5.12 (d, J = 7.7 Hz, J = 6.8 Hz, 2H, 1'H); 4.23 (m, 6H, 2'H, 3'H, 4'H); 3.80, 3.79 (s, 12H, OCH_3); 3.54 (m, 4H, 5'H); 1.17 (m, 12H, $CH(CH_3)_2$); 0.82 (s, 18H, $SiC(CH_3)_3$); –0.05 (s, 6H, $SiCH_3$); –0.17 (s, 6H, $SiCH_3$); ^{31}P-NMR: γ [ppm] (162 MHz, $CDCl_3$): 150.70 and 148.85 (ratio 1:3.8); ESI(+): m/z 863.6 ($[M+H]^+$);

1.2 Oligonucleotide synthesis

The RNA oligomers were synthesized on a Synthesizer D300+ (*Eppendorf*) by phosphoamidite chemistry, with coupling time for modified monomers of 12 min[4]. The fully protected dodecamers were cleaved from the controlled-pore-glass (CPG) support with 1:3 ethanol:NH_3 solution at 40 °C over 24 h. The 2'-silyl groups were deprotected with triethylamine, N-methylpyrrolidinone, and $NEt_3 \cdot 3HF$ mixture for 1.5 h at 65 °C[5]. The crude RNA oligomer was precipitated with BuOH at -80 °C over 30 min, and the fully deprotected RNA was purified by means of anion exchange HPLC (*NucleoPac*-PA-100). The pure oligomer was subsequently desalted (Sephadex-G25). All ribonucleosides were characterized by MALDI-TOF-MS. The masses obtained were in good agreement with the calculated ones.

1.3 UV melting curves

UV melting profiles of the RNA duplexes were recorded in a phosphate buffer containing NaCl (140 mM, pH 7.0) with oligonucleotide concentrations of 2 µM for each strand at wavelength 260 nm[6]. Each melting curve was determined twice.

The temperature range was 0 °C – 65 °C with a heating rate of 0.5 °C. The thermodynamic data were extracted from the melting curves by means of a two state model for the transition from duplex to single strands[7].

1.4 CD spectra

CD spectra of RNA duplexes were recorded at 350 – 200 nm with oligonucleotide concentration of 2 µM for each strand in a phosphate buffer containing NaCl (140 mM, pH 7.0). The measurements were performed at 10 °C to ensure that only duplex RNA was present.

1.5 $\log P$ values

The nucleosides **1** – **5** were dissolved in water, using concentrations that led to an absorbance of 0.8 – 1.2. 2 ml of the solution were mixed with the same volume of octan-1-ol for 10 minutes. The sample was then centrifuged with 10 000 rpm for 10 minutes, and the two phases were seperated. The absorbance of the water and octan-1-ol phases were measured. The partition coefficient was finally calculated according to $\log P = A_{octan-1-ol}/A_{water}$.

2 Computational methods

2.1 Molecular dynamics simulations

The molecular dynamics simulations were performed with AMBER 8[8], using the force field of Cornell et al.[9]. Atomic charges of the fluorinated benzene nucleoside analogues were derived following the RESP[10] method using a 6-31G* basis set. The atomic charges of the corresponding nucleotides were obtained by the fitting procedure described by Cieplak et al.[11]; the charges of the ribose moiety were set to standard values in this case, except for the C1'and H1' atoms. The simulations were started with duplex RNA structures in the canonical A form. The starting structures were generated using the nucgen program of the AMBER suite. The systems were neutralized with Na^+ ions, which were placed at the minima of the electrostatic potential at the surface of the RNA backbone, and solvated in an octahedral box by TIP3P water. The distance between the solute and the boundary of the box was at least 11 Å in all cases. The systems were minimized over 50 steps of steepest descent followed by 450 steps of conjugate gradient, while holding the solute fixed by harmonic restraints with force constants of 25 kcal mol^{-1} Å^{-2} and 5 kcal mol^{-1} Å^{-2}, respectively. The systems were heated from 100 K to 300 K during an MD run under constant NV conditions for 50 ps, whereby the solute was restrained by harmonic restraints with a force constant of 5 kcal mol^{-1} Å^{-2}. Subsequently, the density of the simulation boxes was adjusted under constant NPT conditions for 50 ps, still applying harmonic restraints to the

solute. The harmonic restraints were then removed in steps of 1 kcal mol^{-1} Å^{-2} under constant NVT conditions during additional 50 ps. The equilibration phase was completed by another 100 ps of NVT simulation without any restraints. All MD simulations were performed under periodic boundary conditions, and the particle mesh Ewald method (PME)[12,13] with an 8 Å direct space cut-off and a fourth order B-spline interpolation was applied to treat long-range interactions. The simulations were run with time steps of 2 fs. Bond lengths involving bonds to hydrogen atoms were fixed by using the SHAKE[14,15] algorithm, except for those bonds involving hydrogen atoms with a coupled potential in the perturbed MD runs (see below). The unperturbed duplex simulations were carried on for 10 ns. To keep the temperature constant, the systems were coupled to an external heat bath using the method of Berendsen et al.[16] with a time constant of 2 ps. In the perturbed MD simulations for the free energy calculations (see below), the first 300 ps of the first step and the first 200 ps of all other steps, respectively, were treated as equilibration periods. The following 500 ps at each step were used for the free energy calculations. In the simulations of the single strands, 2 ns were used for the equilibration of the first step instead. Single strand structures were generated from double strand ones as starting structures in these cases.

2.2 Free energy calculations

Relative binding free energies were calculated using the thermodynamic integration (TI) method[17,18]:

$$G_{\mathrm{F}} - G_{\mathrm{H}} = \Delta G \quad = \quad \int_0^1 \left\langle \frac{\mathrm{d}V\left(\mathbf{r},\lambda\right)}{\mathrm{d}\lambda} \right\rangle_{\lambda} \mathrm{d}\lambda \tag{1}$$

The indexes F and H represent the higher and the lower fluorinated state, respectively. G_{F} and G_{H} are the corresponding free energies. V is the potential energy and λ a continuous linear coupling parameter between zero and one that combines the potentials of the higher and lower fluorinated states to a mixed potential of an intermediate state:

$$V\left(\mathbf{r},\lambda\right) \quad = \quad \left(1-\lambda\right) V\left(\mathbf{r},0\right) + \lambda V\left(\mathbf{r},1\right) \tag{2}$$

The angle brackets in equation (1) are denoting ensemble averages over state λ. In the present study MD simulations with the mixed potential at five discrete steps were performed to calculate $\left\langle \mathrm{d}V/\mathrm{d}\lambda \right\rangle_{\lambda}$. The values for λ were chosen such that a Gaussian quadrature could be used to integrate over all λ steps. The integration yields the free energy difference between two differently fluorinated states.

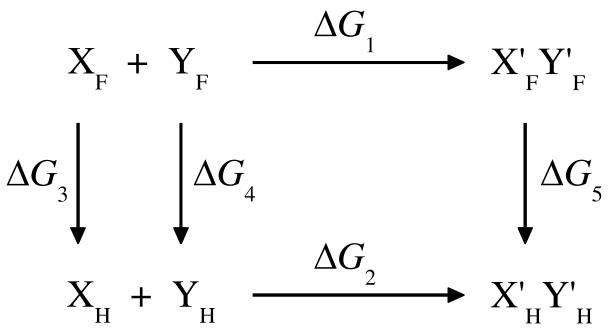

Figure S1. Thermodynamic cycle applied for calculating relativ binding free energies.

The binding free energy differences were then calculated according to the thermodynamic cycle in Figure S1. Here X and Y represent the two RNA single strands, each containing a fluorobenzene nucleotide at complementary positions, and X'Y' the RNA double strand with indexes F and H again indicating a higher and lower fluorinated state, respectively. The change in stability (relative binding free energy) $\Delta\Delta G$ (= ΔG_1 - ΔG_2) is obtained according to the thermodynamic cycle as:

$$\Delta\Delta G \;=\; \Delta G_1 - \Delta G_2 = \Delta G_3 + \Delta G_4 - \Delta G_5 \tag{3}$$

As the reduction of the radius of an atom causes a smaller perturbation in its environment than the expansion, the TI simulations were always performed by converting fluorine to hydrogen. Thus, relative binding free energies were calculated for the following transitions: 4-fluorobenzene to benzene, 2,4-difluorobenzene to 4-fluorobenzene, 2,4,5-trifluorobenzene to 2,4-difluorobenzene, and 2,3,4,5-tetrafluorobenzene to 2,4,5-trifluorobenzene. As a control, 2,3,4,5-tetrafluorobenzene was converted to benzene. The change in stability should be equal to the sum of the above calculated relative binding free energies for each step.

2.3 Free energy decomposition

A decomposition into contributions from individual nucleotides, water molecules, and ions to $\mathrm{d}V/\mathrm{d}\lambda$ was implemented into the TI module of the AMBER 8 suite. Bonded interactions are decomposed equally between the atoms involved, as are van der Waals and electrostatic interactions[19], following the arbitrary but intuitive choice that each atom in an interaction contributes equally to that interaction[20].

$$\begin{aligned}\frac{\mathrm{d}V_i}{\mathrm{d}\lambda} \;=\; & \sum_{j\in 1\ldots 2}\left(\frac{1}{2}\frac{\mathrm{d}V_{i,j}^{\mathrm{bond}}}{\mathrm{d}\lambda} + \sum_{k\in 1\ldots 3}\left(\frac{1}{3}\frac{\mathrm{d}V_{i,j,k}^{\mathrm{angle}}}{\mathrm{d}\lambda} + \sum_{l\in 1\ldots 4}\frac{1}{4}\frac{\mathrm{d}V_{i,j,k,l}^{\mathrm{dihedral}}}{\mathrm{d}\lambda}\right)\right) + \\ & + \sum_{j\neq i}\left(\frac{1}{2}\frac{\mathrm{d}V_{i,j}^{\mathrm{direct}}}{\mathrm{d}\lambda}\right) + \frac{\mathrm{d}V_i^{\mathrm{indirect}}}{\mathrm{d}\lambda}\end{aligned} \tag{4}$$

With $V_{i,j}^{\text{bond}}$ being the contribution of the bond length to a covalently bound atom j, $V_{i,j,k}^{\text{angle}}$ the contribution of the bond angle between the covalently bound atoms i, j, and k, and $V_{i,j,k,l}^{\text{dihedral}}$ the contribution of the dihedral angle, spanned by the four covalently bound atoms i, j, k and l. $V_{i,j}^{\text{direct}}$ is the contribution from short range non-bonded interactions calculated in direct space, while V_i^{indirect} is the contribution from long range non-bonded interactions calculated in Fourier space. Contributions to the relative binding free energy per atom are then obtained by $\Delta G_i = \int_0^1 \langle \mathrm{d}V_i/\mathrm{d}\lambda \rangle_\lambda \, \mathrm{d}\lambda$. The contribution of a single nucleotide is obtained by summing over all atom contributions belonging to this nucleotide. Contributions by individual water molecules and counter ions are summed to one overall contribution of the solvent, which reflects a (de-)solvation contribution to the relative binding free energy.

2.4 Calculation of solvent accessible surface area

The calculation of the solvent accessible surface area ($SASA$) was done with the program MS[21]. We used a surface point density of 10.0 Å^{-2}, a probe radius of 1.4 Å, and Bondi radii[22]. The last 7 ns of the unperturbed MD trajectories of the RNA duplexes were used for the calculations, with snapshots sampled every 10 ps. We calculated $SASA$ for the duplex structures and the single strands extracted from these duplexes. The $SASA$ burried upon duplex formation is calculated as the difference between the $SASA$ of the duplex minus the sum of the $SASA$ of the single strand RNAs. The relative buried solvent accessible surface area $\Delta\Delta SASA$ was then calculated as the difference in $\Delta SASA$ for two fluorinated nucleotides. Atomic contributions of the fluorine atoms to $\Delta SASA$ and $\Delta\Delta SASA$ were likewise calculated. In addition to the mean, the standard error of the mean is given as a measure for the precision of the calculations.

2.5 Analysis of C-F$\cdots$H-C interactions

To identify C-F$\cdots$H-C interactions, we analyzed the last 7 ns of the unperturbed MD trajectories of the RNA duplexes in terms of F$\cdots$H distances and C-H$\cdots$F angles with snapshots sampled every 10 ps. To calculate occupancies for the observed interactions, i. e., ratios of times where the interaction is present relative to the total simulation time, we considered an interaction to be present if its F$\cdots$H distance was lower than or equal 3.0 Å. The analyses were done for all possible C-H/F-C pairs between the fluorinated benzene nucleotide self pairs.

2.6 Analysis of helical parameters

Helical parameters of the structures of the last 7 ns of the unperturbed MD trajectories of the RNA duplexes were analysed using the program Curves[23].

3 Results

3.1 Melting temperatures and binding free energies

Melting temperatures and binding free energies extracted from UV melting profiles for duplex RNA with nucleotide analogue self-pairs are shown in Table S1.

Table S1. Melting temperatures and free energies extracted from the UV melting profiles.

Self pair	Melting temperature[a]	ΔG[b]
benzene	26.4 ± 0.5	8.5 ± 0.2
4-fluorobenzene	32.5 ± 0.7	10.2 ± 0.2
2,4-difluorobenzene	35.2 ± 0.7	11.6 ± 0.2
2,4,5-trifluorobenzene	35.8 ± 0.7	11.8 ± 0.2
2,3,4,5-tetrafluorobenzene	38.0 ± 0.8	12.2 ± 0.2

[a] In °C. [b] In kcal mol^{-1}.

3.2 CD spectra

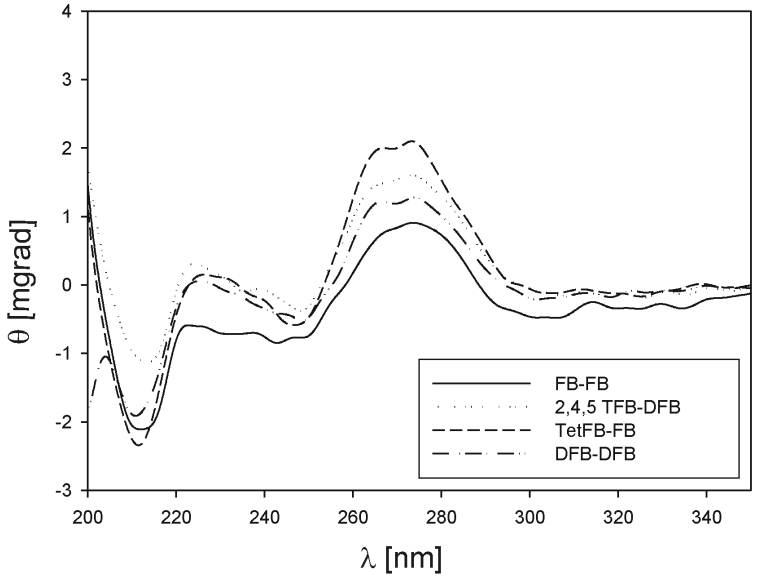

Figure S2. CD spectra of representative combinations of modified RNA duplexes (5´-CUU UUC XUU CUU-3´ paired with 3´-GAA AAG YAA GAA-5´), where X and Y are FB = 4-fluorobenzene (2), DFB = 2,4-difluorobenzene (3), 2,4,5-TFB = 2,4,5-trifluorobenzene (4), TetFB = 2,3,4,5-tetrafluorobenzene (5).

3.3 Free energy calculations

In Figure S3 $\langle \mathrm{d}V/\mathrm{d}\lambda \rangle_\lambda$ is depicted as a function of the coupling parameter λ for each transition, applying the free energy decomposition scheme. The values of the TI simulations of the two single strands are added, resulting in a value for the unbound state. The nearly linear dependence of $\langle \mathrm{d}V/\mathrm{d}\lambda \rangle_\lambda$ with respect to λ enables very precise calculations of $\Delta\Delta G$, even if only five λ steps are used. The correlation between the measured and calculated relative binding free energies is shown in Figure S4. For both the free energy calculations with and without decomposition, very good correlations with the experimentally determined values are obtained ($R^2 > 0.96$).

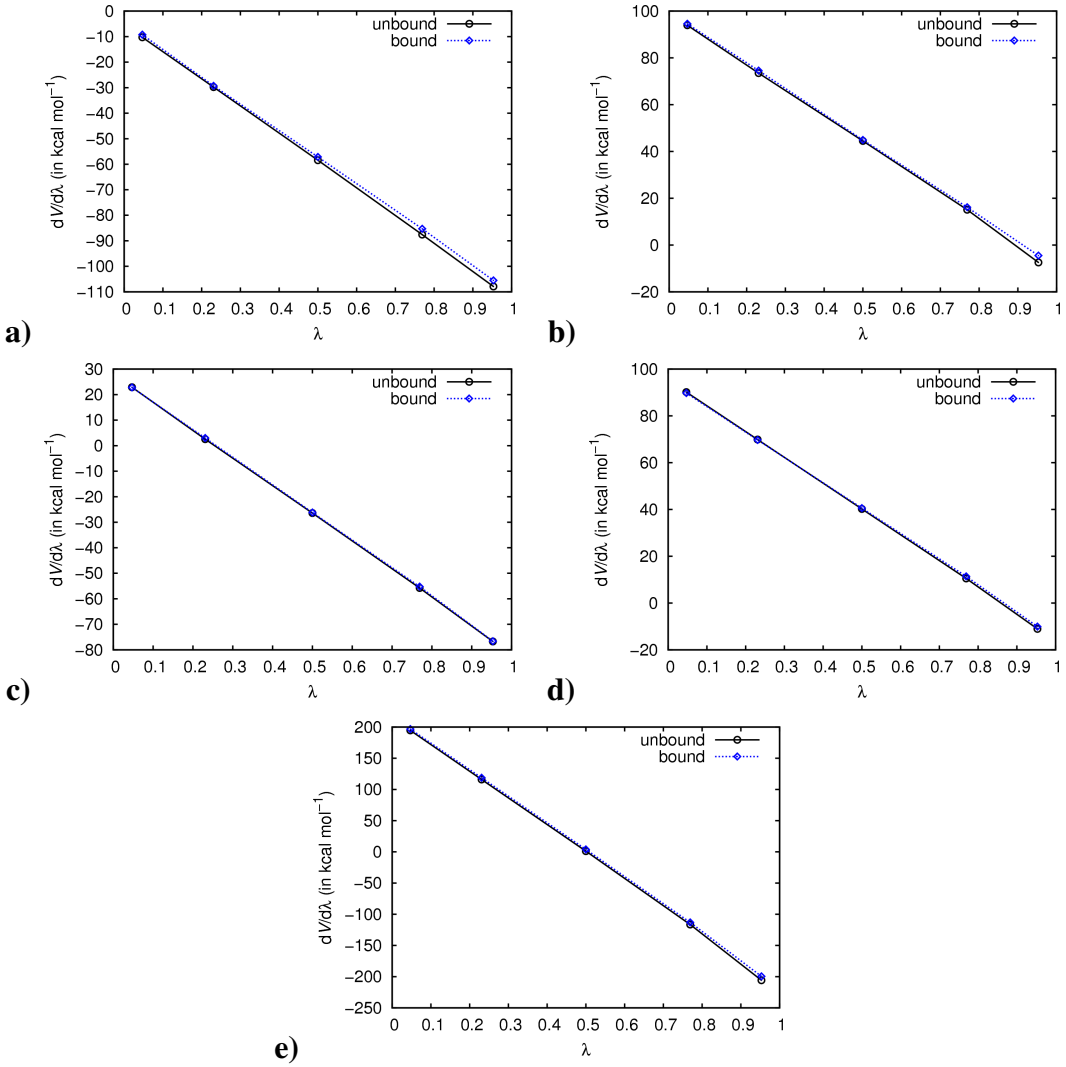

Figure S3. $\langle \mathrm{d}V/\mathrm{d}\lambda \rangle_\lambda$ as a function of λ for the free energy simulations of the unbound and bound state of each system: **a)** benzene → 4-fluorobenzene, **b)** 4-fluorobenzene → 2,4-difluorobenzene, **c)** 2,4-difluorobenzene → 2,4,5-trifluorobenzene, **d)** 2,4,5-trifluorobenzene → 2,3,4,5-tetrafluorobenzene and **e)** benzene → 2,3,4,5-tetrafluorobenzene.

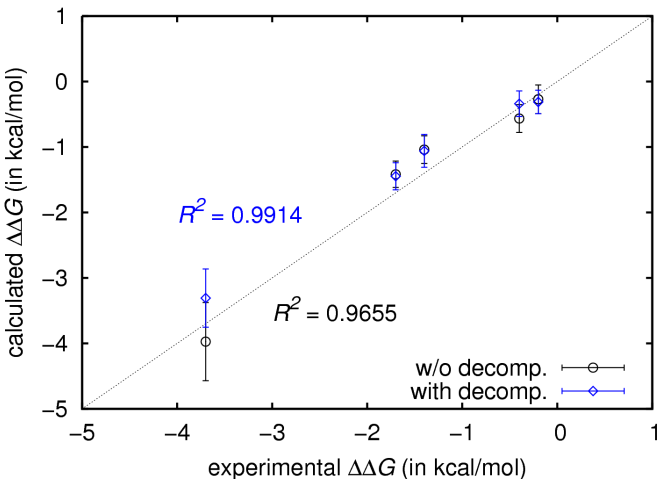

Figure S4. Calculated $\Delta\Delta G$ versus experimental $\Delta\Delta G$. The dashed line shows the line of ideal correlation.

3.4 Solvent accessible surface area

The solvent accessible surface area parts contributed by fluorine atoms that are buried upon double strand formation $\Delta SASA_F$ are given in Table S2.

Table S2. Calculated fluorine contributions to the buried solvent accessible surface area.

Self pair	$\Delta SASA_F$[a]
4-fluorobenzene	13.7 ± 0.3
2,4-difluorobenzene	15.9 ± 0.3
2,4,5-trifluorobenzene	17.4 ± 0.4
2,3,4,5-tetrafluorobenzene	32.8 ± 0.6

[a] In Å.

3.5 Analysis of C-F$\cdots$H-C interactions

In all three systems in which intermolecular C-F$\cdots$H-C interactions can occur (self-paired 4-fluorobenzene, 2,4-difluorobenzene, and 2,4,5-trifluorobenzene), the pairing of C3-H3 in the X-strand with F4-C4 in the Y-strand shows the shortest F$\cdots$H distance. Scatterplots of the F$\cdots$H distance with respect to the C-H$\cdots$F angle are shown in Figure S5. Shorter F$\cdots$H distances correlate with larger C-H$\cdots$F angles in all cases, which indicates a dipole-dipole interaction. Compared to self-paired 4-fluorobenzene, shorter distances and larger angles occur more frequently in self-paired 2,4-difluorobenzene and 2,4,5-trifluorobenzene.

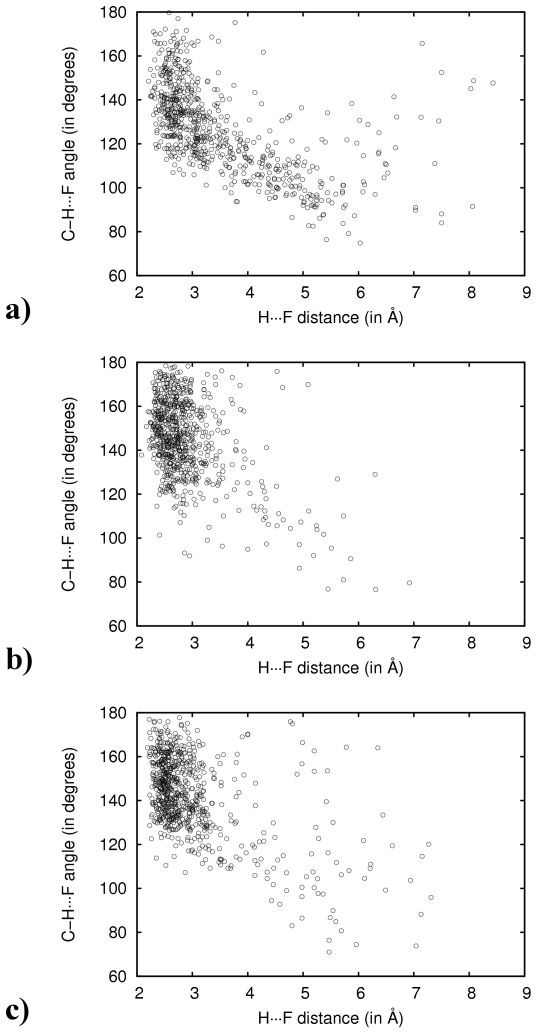

Figure S5. Scatterplots of the C-H$\cdots$F angle versus the H$\cdots$F distance for the interactions between C3-H3 in the X-strand with F4 in the Y-strand for the systems with **a)** 4-fluorobenzene, **b)** 2,4-difluorobenzene and **c)** 2,3,4-trifluorobenzene self pairs.

References

[1] J. Parsch, J. W. Engels, *Helv. Chim. Acta* **2000**, *83*, 1791–1808.

[2] J. Parsch, J. W. Engels, *J. Am. Chem. Soc.* **2002**, *124*, 5664–5672.

[3] A. Živković, Dissertation, University of Frankfurt am Main (Germany), **2005**.

[4] M. H. Lyttle, P. B. Wright, N. D. Sinha, J. D. Bain, A. R. Chamberlin, *J. Org. Chem.* **1991**, *56*(15), 4608–4615.

[5] F. E. Wincott, N. Usman, *Methods in Mol. Biology* **2000**, *74*, 59.

[6] M. Schweitzer, J. W. Engels, *Antisense: From Technology to Therapy*, Blackwell Science, Cambridge, MA, Bd. 6, **1997**, 78–103.

[7] L. A. Marky, K. J. Breslauer, *Biopolymers* **1987**, *26*(9), 1601–1620.

[8] D. A. Case, T. E. Cheatham, III., T. A. Darden, H. Gohlke, R. Luo, K. M. Merz, A. Onufriev, C. Simmerling, B. Wang, R. J. Woods, *J. Comput. Chem.* **2005**, *26*(16), 1668–1688.

[9] W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell, P. A. Kollman, *J. Am. Chem. Soc.* **1995**, *117*, 5179–5197.

[10] C. I. Bayly, P. Cieplak, W. D. Cornell, P. A. Cornell, *J. Phys. Chem.* **1993**, *97*, 10269–10280.

[11] P. Cieplak, W. D. Cornell, C. Bayly, P. A. Kollman, *J. Comput. Chem.* **1995**, *16*(11), 1357–1377.

[12] T. Darden, D. York, L. Pedersen, *J. Chem. Phys.* **1993**, *98*(12), 10089–10092.

[13] T. E.Cheatham, III., J. L. Miller, T. Fox, T. A. Darden, P. A. Kollman, *J. Am. Chem. Soc.* **1995**, *117*(14), 4193–4194.

[14] J.-P. Ryckaert, G. Ciccotti, H. J. C. Berendsen, *J. Comput. Phys.* **1977**, *23*(3), 327–341.

[15] S. Miyamoto, P. A. Kollman, *J. Comput. Chem.* **1992**, *13*(8), 952–962.

[16] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola, J. R. Haak, *J. Chem. Phys.* **1984**, *81*(8), 3684–3690.

[17] J. G. Kirkwood, *J. Chem. Phys.* **1935**, *3*(5), 300–313.

[18] D. L. Beveridge, F. M. DiCapua, *Annu. Rev. Biophys. Biophys. Chem.* **1989**, *18*, 431–492.

[19] H. Gohlke, C. Kiel, D. A. Case, *J. Mol. Biol.* **2003**, *330*(4), 891–913.

[20] D. A. Pearlman, D. A. Case, J. W. Caldwell, W. S. Ross, T. E. Cheatham, III., S. DeBolt, D. Ferguson, G. Seibel, P. Kollman, *Comput. Phys. Commun.* **1995**, *91*(1–3), 1–41.

[21] M. L. Connolly, *Science* **1983**, *221*(4612), 709–713.

[22] A. Bondi, *J. Phys. Chem.* **1964**, *68*, 441–451.

[23] R. Lavery, H. Sklenar, *J. Biomol. Struct. Dyn.* **1988**, *6*, 63–91.