1

Synthesis and Properties of Fluorinated Nucleobases in DNA and RNA

Holger Gohlke,* Jelena Bozilovic* and Joachim W. Engels*

1.1 Introduction

The stability of nucleic acid structures is predominantly governed by hydrogen bonding, base stacking and solvation. To probe these interactions, a common approach is to replace the native bases adenine (A), uracil (U)/thymidine (T), guanosine (G) and cytosine (C) with analogues in which functional groups are added, deleted, blocked or rearranged. The size and shape of the analogues should be preserved as closely as possible to native bases. Such ‘nonpolar nucleoside isosteres’ (NNIs) then allow detection predominant forces within nucleic acid structures without introducing steric effects. For DNA, this concept was introduced by Kool and coworkers in 1994 (Schweitzer and Kool, 1994). Initially, these molecules were intended to be used as probes of the importance of hydrogen bonding and base stacking in the formation of stable DNA duplex structures (Schweitzer and Kool, 1994; Kool and Sintim, 2006). In the context of DNA replication, it was later concluded by these authors that steric
effects rather than hydrogen bonding was the chief explanation for replication fidelity (Kool, 1998; Kool and Sintim, 2006). In 1999, the concept of NNI was introduced into the RNA world by Engels and coworkers, initially using fluorinated benzenes and benzimidazoles as pyrimidine and purine base analogues, respectively (Parsch and Engels, 1999, 2000). Syntheses and crystallographic studies of these and other NNIs as well as thermodynamic analyses and computer simulations of model RNAs incorporating them are reviewed here. These combined studies have proved invaluable for probing the physical forces that govern the stability of RNA and shedding light on the role of fluorine in molecular recognition. As a biological application, the incorporation of fluorinated NNI into ribozymes and siRNA is finally described.

1.2 Fluorine in Molecular Recognition

Substituting hydrogen by fluorine in organic compounds influences a variety of the molecule's properties. In medicinal chemistry, fluorine substitution has long been known as a means of enhancing metabolic stability, modifying chemical reactivity and conformational equilibria and improving transportation and absorption characteristics of pharmaceuticals (Muller et al., 2007; Begue and Bonnet-Delpont, 2008; Hagmann, 2008). In contrast, the role of ‘organic fluorine’ in influencing molecular recognition properties, i.e. specific bonding between two or more molecules through non-covalent interactions, is much less understood. Here, we focus on influences that are particularly important in the context of fluorinated NNIs.

The properties of the C–F bond provide a starting point for appreciating some of these influences. Replacing F for H is considered the most conservative substitution for hydrogen on steric grounds, although a fluorine atom is closer in size (and bond length) to an oxygen than hydrogen (O’Hagan, 2008). Hydrogen and fluorine are also quite different regarding their electronic influences. The high electronegativity of fluorine imparts a less covalent and more electrostatic character to a highly polarized C–F bond, allowing interactions between the C–F bond dipole and other dipoles in close proximity (O’Hagan, 2008). In turn, fluorine’s three lone pairs are held tightly, as manifested by the atom’s high ionization potential.
and low polarizability, and so are reluctant to get involved in resonance or act as hydrogen bonding acceptors.

As an immediate consequence, fluorine substitution leads to the seemingly orthogonal effects of increasing local polarity and molecular hydrophobicity (Guerra and Bickelhaupt, 2003; Biffinger et al., 2004; DiMagno and Sun, 2006): as electrostatic and, in particular, time-dependent interactions of C–F bonds are of minimal importance in polar heteroatom solvents, C–F dipoles and lone pairs are poorly solvated in water. Removing this group from water is thus energetically favourable. Accordingly, for a set of 293 pairs of compounds where a single H/F exchange had been performed, an increase of the log $D$ value by on average 0.25 log units upon fluorine substitution was observed (Bohm et al., 2004). Notably, for a series of singly and multiply fluoro-substituted benzenes, no impact of the molecular dipole moment on the partition coefficient was found, suggesting that surface area arguments suffice to explain hydrophobicity trends (DiMagno and Sun, 2006).

Due to the electron-withdrawing effect of fluorine, fluorination has an important indirect impact on interactions of an aromatic ring to which it is attached. On going from benzene to hexafluorobenzene, the quadrupole moment of the molecules inverts (Hernandeztrujillo and Vela, 1996), favouring eclipsed face-to-face π-stacks in crystallized 1:1 mixtures (Patrick and Prosser, 1960). Similarly, using the dangling-end method for determining the energetics of aromatic π-stacking of fluorinated benzene analogues of DNA bases, a large stabilizing interaction is observed with a 2,3,4,5-tetrafluorophenyl dangling nucleotide. Yet, the overall correlation between the number of fluorine groups and the stabilization is only weak (Lai et al., 2003). The σ-inductive effect of fluorine also affects neighbouring functions. With respect to fluorinated NNIs, the influence on the acidity of neighbouring C–H groups is of prime importance: the C–H group is known to be a hydrogen-bond donor (Desiraju, 1997) and a C–H···X interaction is strengthened by a more positively polarized C–H group (Thalladi et al., 1998). A particularly strong polarization is expected for a hydrogen in bis-ortho position to two fluorine atoms, e.g. in 1,3-difluorobenzene or 1,3,5-trifluorobenzene (Razgulin and Mecozzi, 2006). Not surprisingly, in the crystal structure of 1,3,5-trifluorobenzene, each H and F atom is involved in the formation of two C–H···F interactions...
(Thalladi et al., 1998), leading to a close similarity to the classical structure of 1,3,5-triazine (Coppens, 1967). Obviously, the C–H⋯F interaction is favoured by multipolar interactions between the C–H / F–C bond dipoles, as also demonstrated by numerous O=C⋯F–C contacts found in the Cambridge Structural Database (Olsen et al., 2003).

In the presence of competing heteroatom acceptors, covalently bound fluorine hardly ever acts as acceptor for available Bronsted acidic sites, owing to its low proton affinity and weak polarizability (Murrayrust et al., 1981; Dunitz and Taylor, 1997; Dunitz, 2004; Kool and Sintim, 2006). Attractive C–F⋯H–X dipolar interactions have been described, however, for well-structured molecular environments in which heteroatom acceptors are excluded, such as the thrombin active site (Olsen et al., 2003; Bohm et al., 2004) or engineered crystals (Desiraju, 2002; Reichenbacher et al., 2005).

1.3 Synthesis of Fluoro-substituted Benzenes, Benzimidazoles and Indoles, and their Incorporation into Model RNA

1.3.1 Chemical syntheses of fluoro substituted benzenes, benzimidazoles and indoles

When we started to synthesize fluorobenzene nucleosides 1–4 (Fig. 1.1) some 10 years ago we were able to directly introduce the lithiated fluorobenzenes to the protected ribonolactone.

The method is based on a similar procedure introduced by Krohn et al. (1992), where a bromo-fluorobenzene is added to a benzylated ribonolactone. The resulting lactol is reduced and yields directly the pure β-ribonucleoside in high yield.

Figure 1.1. Initial set of fluorobenzene ribosides.
In Fig. 1.2, the synthesis of 2,4-difluorobenzene riboside as phosphoramidite building block is shown as an example (Parsch and Engels, 2002). The unprotected riboside is tritylated with dimethoxytritylchloride in pyridine with high regioselectivity for the 5'-position, followed by silylation with tert-butyldimethylsilyl chloride for protecting the 2'- or 3'-positions. After a chromatographic separation, the 2'-silylated nucleoside is phosphitylated with the chloro-cyanoethyl-diisopropylphosphoramidite. The overall yields for the four nucleobases prepared in an analogous way range from 10 to 20%.

In Fig. 1.3, as an example, the synthesis of 1'-deoxy-1'-(4-fluorobenzimidazol-1-yl)-β-D-ribofuranose 21 is shown, following the glycosylation procedure of Vorbruggen (Vorbruggen and Hofle, 1981). Refluxing two
equivalents of 4-fluorobenzimidazole 17 with N,O-bis(trimethylsilyl) acetamide and subsequent reaction of the persilylated base with one equivalent of 1,2,3,5-tetra-O-acetyl-β-D-ribofuranose 18 in the presence of the Lewis acid trimethylsilyl-trifluoromethanesulfonate afforded the desired 2',3',5'-tri-O-acetyl-1'-deoxy-1'-(4-fluorobenzimidazol-1-yl)-β-D-ribofuranose 19 in 56% yield. Using microwave reaction conditions we could improve the yields (Nikolaus et al., 2007). As a by-product, a bis-glycosylated benzimidazole was obtained. Deprotection of the acetylated nucleoside 19 furnished 1'-deoxy-1'-(4-fluorobenzimidazol-1-yl)-β-D-ribofuranose 21 in 90% yield (Bats et al., 2000; Parsch and Engels, 2000, 2001). The phosphoramidites were obtained in an analogous way as described above and the yields ranged from 10 to 15% (Parsch and Engels, 2000, 2002).

Syntheses of the fluorobenzimidazole ribosides 21 and 39–41 started with the appropriately substituted fluoroacetanilides, followed by nitration

![Figure 1.3. Synthesis of a fluorobenzimidazole nucleoside phosphoramidite.](image-url)
and ring closure with formic acid (Parsch and Engels, 2000, 2002). The trifluoromethylbenzimidazoles 42–45 were analogously prepared starting from the trifluoromethyl acetanilides (Moore et al., 2004). The log P values measured in octanol/water for the mono-substituted benzimidazoles are identical and the orientation of the fluoro-substituent does not seem to have a major influence here. The difluorobenzimidazole riboside 41 is significantly more lipophilic (log P = 4.2) (Table 1.1).

Since the indole structure is present in a variety of natural compounds, such as amino acids and alkaloids, we decided to synthesize and evaluate a similar pattern of substitution on fluoroindoles. Furthermore, the charge distribution and dipole moments between fluoroindole and fluorobenzimidazole compounds differ significantly, which we also expected to give rise to differences in stability of RNA oligonucleotides containing these building blocks.

To the best of our knowledge, ribofluoro-indole compounds had not been synthesized and evaluated before. We reported the successful syntheses of fluoroindole building blocks and the X-ray crystal structures of all synthesized fluoroindole ribonucleosides 23–27 (Fig. 1.4) (Bozilovic and Engels, 2007).

The synthesis of fluoroindoles, which are not commercially available, was achieved very efficiently in a four-step procedure as shown in Fig. 1.5 for the 4,6-difluoroindol-β-D-ribofuranoside 27. For the 4,6-difluoroindol

### Table 1.1. Partition coefficients (log P values) of nucleosides containing fluoro substituted nucleobase analogues.

<table>
<thead>
<tr>
<th>Nucleoside</th>
<th>Partition coefficient</th>
<th>Nucleoside</th>
<th>Partition coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>B</td>
<td>1.050</td>
<td>4,6DFBI</td>
<td>4.240</td>
</tr>
<tr>
<td>4FB</td>
<td>1.497</td>
<td>4TFM</td>
<td>6.762</td>
</tr>
<tr>
<td>2,4DFB</td>
<td>1.695</td>
<td>5TFM</td>
<td>6.757</td>
</tr>
<tr>
<td>2,4,6TFB</td>
<td>1.398</td>
<td>6TFM</td>
<td>6.688</td>
</tr>
<tr>
<td>2,4,5TFB</td>
<td>1.825</td>
<td>I</td>
<td>0.72</td>
</tr>
<tr>
<td>PFB</td>
<td>2.156</td>
<td>4FI</td>
<td>2.781</td>
</tr>
<tr>
<td>BI</td>
<td>0.149</td>
<td>5FI</td>
<td>2.252</td>
</tr>
<tr>
<td>4FBI</td>
<td>1.789</td>
<td>6FI</td>
<td>4.420</td>
</tr>
<tr>
<td>5FBI</td>
<td>1.802</td>
<td>7FI</td>
<td>2.383</td>
</tr>
<tr>
<td>6FBI</td>
<td>1.799</td>
<td>4,6DFI</td>
<td>5.540</td>
</tr>
</tbody>
</table>

1. Synthesis and Properties of Fluorinated Nucleobases in DNA and RNA
2. Fluorine in Pharmaceutical and Medicinal Chemistry
3. 1st Reading
30, methyl azidoacetate, an intermediate in indole synthesis, was synthesized by a literature procedure (Da Rosa et al., 2003). A prolonged reaction time and lower reaction temperature improved this procedure (Bozilovic and Engels, 2007). The hydrolysis of the methyl ester is followed by a decarboxylation, where the main problem lies in the very high temperature at which the reaction occurs. Prolonging the reaction time does not lead to higher yields (Bozilovic and Engels, 2007). Since for the indole moiety direct glycosylation with the ribose failed, synthesis of the deoxyribose and subsequent transformation to the ribose was chosen as an
Figure 1.5. Summary of fluoro base analogues comprising fluorobenzenes, fluorobenzimidazoles, fluoroindoles, and trifluoromethyl benzimidazoles.
alternative. Although this is a longer procedure, the individual steps could be optimized. Direct S_N2 substitution of the indole anion with the chlorosugar 30 yielded 32 quantitatively. After deprotection, the 5'-OH is protected again with tertiary butyl dimethylsilyl chloride (TBDPSi–Cl) and 3'-OH with mesylchloride (MsCl). Deprotection of 5'-OH and elimination of mesylate gave 35 in one step, which on dihydroxylation via osmium tetroxide resulted in the formation of indole riboside 27 in an overall yield of 52% starting from 29. For the RNA building block the phosphoramidite was prepared as outlined above (Fig. 1.2). The synthesis of the indole ribosides 23–26 followed the same procedure. The total set of fluoro-modified benzenes, benzimidazoles and indoles is outlined in Fig. 1.5.

1.3.2 Synthesis of 12-mer RNA duplexes that incorporate fluoronucleosides

The phosphoramidites 11 and 22 as well as the phosphoramidites of all other fluorobenzenes, fluorobenzimidazoles and fluoroindoles were incorporated into a defined 12-mer RNA duplex to investigate their influence on RNA stability. We chose the purine rich sequences 5'-CUUUUCXUUCUU-3 paired with 3'-GAAAAGYAAGAA-5, where the central bases X or Y contained modified fluoronucleosides. This yielded either mono-modified duplex RNA, where a fluorinated NNI is paired against a natural base, or double-modified duplex RNA containing fluorinated NNIs paired against each other. Solid phase RNA synthesis followed the Caruthers DNA/RNA cycle (Fig. 1.6) (Caruthers, 1985).

In short, a starting riboside, attached to controlled-pore glass (CPG) support on the 3'-end, is stepwise extended via tetrazole activation with an appropriately protected amidite of A, C, G, or U. In the case of X or Y, the fluorinated NNI is added and the sequence further extended until the 5'-end. After successful additions, which are monitored by the trityl colour, the full size oligoribonucleotide is liberated from its protecting groups and the crude product is purified. In most cases high performance liquid chromatography on ion exchange columns or sometimes reversed-phased material gives sufficiently pure oligoribonucleotides, which are characterized by mass spectrometry. Notably, the CD spectra of the RNA
duplexes with fluorinated NNI showed the typical curves for an A-type helix (Fig. 1.7), indicating that the structure of the duplex RNA is not disturbed by incorporation of the modified nucleosides.

1.3.3 RNA melting studies and thermodynamic data

Thermodynamic stabilities of the modified RNA duplexes were determined by thermal denaturation as monitored by UV absorbance in a phosphate buffer (10 mM Na₂HPO₄, 10 mM NaH₂PO₄, pH 7) containing NaCl (140 mM). Thermodynamic data were extracted from the melting curves by means of a two-state model for the transition from duplex to single strands.
Melting temperatures ($T_m$) and free energies of duplex stability ($\Delta G$) for all fluorobenzenes, -benzimidazoles and -indoles paired against all four natural bases are summarized in Table 1.2. Thermodynamic analysis of the $T_m$ furthermore yielded enthalpic ($\Delta H$) and entropic ($T \Delta S$) contributions to the duplex stability (data not shown) (Parsch and Engels, 2002).

Two results stand out from these studies. First, compared to the natural U:A base pair ($T_m = 37.8^\circ C$) or the wobble base pair U:G ($T_m = 38.6^\circ C$), pairing of fluoronucleosides against natural bases in general decreases RNA duplex stability, as demonstrated by $T_m$ values that are lower by 2–15$^\circ C$. The lowest $T_m$ were measured for 36 and 47 due to steric effects resulting from a bis-ortho fluoro substitution. Conversely, pairing 40 or 25 against A almost restored the duplex stability to that of the natural U:A base pair, with $T_m$ values of 34.0$^\circ C$ and 37.1$^\circ C$, respectively. Second and more strikingly, several of these fluoro-benzenes, -benzimidazoles and -indoles showed a lack of discrimination against the natural bases, thus acting as universal bases (Loakes, 2001). In the case of 4 and 41,
the \( T_m \) values differ by less than 1°C, yielding these fluoronucleosides as the best universal bases.

In a second series, the stability of RNA duplexes containing fluorobenzene self pairs at positions X and Y were determined (Table 1.3). Not unexpectedly (Parsch and Engels, 2002; Lai and Kool, 2004), the measurements demonstrated that the pairing preference of fluorinated NNI is higher in self-pairs than in pairs with natural bases. Furthermore, the duplex stability increases incrementally with the number of fluorine substituents in the NNI. Surprisingly, this leads to RNA duplex stabilities with self-paired bases 4, 37, and 38 (\( T_m = 35.2, 35.8, 38.0°C \)) that are similar to or exceed that of the natural U:A base pair.

### Table 1.2. Thermodynamic stabilities of model 12mer RNA containing fluorobenzenes (FB), fluorobenzimidazoles (FBI), trifluoromethylbenzimidazoles (TFM) and fluorin- doles (FI) paired with natural bases A, C, G, and U.

<table>
<thead>
<tr>
<th>X</th>
<th>Y=A ( T_m ) [a]</th>
<th>( \Delta G ) [b]</th>
<th>Y=C ( T_m ) [a]</th>
<th>( \Delta G ) [b]</th>
<th>Y=G ( T_m ) [a]</th>
<th>( \Delta G ) [b]</th>
<th>Y=U ( T_m ) [a]</th>
<th>( \Delta G ) [b]</th>
</tr>
</thead>
<tbody>
<tr>
<td>4FB</td>
<td>23.8</td>
<td>7.9</td>
<td>24.1</td>
<td>8.0</td>
<td>24.2</td>
<td>8.0</td>
<td>25.6</td>
<td>8.4</td>
</tr>
<tr>
<td>2,4DFB</td>
<td>27.4</td>
<td>9.0</td>
<td>27.3</td>
<td>8.9</td>
<td>27.6</td>
<td>9.0</td>
<td>27.9</td>
<td>9.1</td>
</tr>
<tr>
<td>2,4,6TFB</td>
<td>23.3</td>
<td>7.8</td>
<td>20.6</td>
<td>7.1</td>
<td>22.8</td>
<td>7.7</td>
<td>22.9</td>
<td>7.9</td>
</tr>
<tr>
<td>2,4,5TFB</td>
<td>25.6</td>
<td>8.4</td>
<td>26.7</td>
<td>8.8</td>
<td>28.7</td>
<td>9.3</td>
<td>27.5</td>
<td>8.9</td>
</tr>
<tr>
<td>PFB</td>
<td>22.9</td>
<td>7.7</td>
<td>21.7</td>
<td>7.2</td>
<td>23.0</td>
<td>7.7</td>
<td>23.1</td>
<td>7.5</td>
</tr>
<tr>
<td>4FBI</td>
<td>28.0</td>
<td>9.1</td>
<td>27.5</td>
<td>8.9</td>
<td>28.7</td>
<td>9.3</td>
<td>28.5</td>
<td>9.2</td>
</tr>
<tr>
<td>5FBI</td>
<td>28.9</td>
<td>9.1</td>
<td>31.0</td>
<td>9.2</td>
<td>31.7</td>
<td>9.3</td>
<td>28.2</td>
<td>8.9</td>
</tr>
<tr>
<td>6FBI</td>
<td>34.0</td>
<td>7.5</td>
<td>31.3</td>
<td>10.1</td>
<td>32.4</td>
<td>10.3</td>
<td>28.6</td>
<td>9.3</td>
</tr>
<tr>
<td>4,6DFBI</td>
<td>28.4</td>
<td>9.2</td>
<td>28.7</td>
<td>9.2</td>
<td>29.4</td>
<td>9.5</td>
<td>29.3</td>
<td>9.5</td>
</tr>
<tr>
<td>4TFM</td>
<td>30.1</td>
<td>11.0</td>
<td>27.4</td>
<td>9.9</td>
<td>29.2</td>
<td>9.7</td>
<td>25.1</td>
<td>9.3</td>
</tr>
<tr>
<td>5TFM</td>
<td>30.2</td>
<td>9.1</td>
<td>28.1</td>
<td>8.8</td>
<td>29.3</td>
<td>8.7</td>
<td>24.1</td>
<td>8.1</td>
</tr>
<tr>
<td>6TFM</td>
<td>30.1</td>
<td>10.9</td>
<td>28.0</td>
<td>8.9</td>
<td>27.1</td>
<td>9.6</td>
<td>24.8</td>
<td>8.2</td>
</tr>
<tr>
<td>4FI</td>
<td>25.1</td>
<td>8.3</td>
<td>24.1</td>
<td>8.0</td>
<td>21.4</td>
<td>7.6</td>
<td>21.0</td>
<td>7.2</td>
</tr>
<tr>
<td>5FI</td>
<td>32.8</td>
<td>9.7</td>
<td>26.8</td>
<td>8.6</td>
<td>29.3</td>
<td>9.3</td>
<td>25.0</td>
<td>8.2</td>
</tr>
<tr>
<td>6FI</td>
<td>37.1</td>
<td>10.6</td>
<td>34.6</td>
<td>15.4</td>
<td>35.5</td>
<td>11.3</td>
<td>31.7</td>
<td>9.7</td>
</tr>
<tr>
<td>7FI</td>
<td>34.4</td>
<td>9.9</td>
<td>24.8</td>
<td>8.2</td>
<td>24.6</td>
<td>8.3</td>
<td>29.8</td>
<td>9.5</td>
</tr>
<tr>
<td>4,6DFI</td>
<td>33.8</td>
<td>8.4</td>
<td>32.0</td>
<td>9.2</td>
<td>30.2</td>
<td>9.4</td>
<td>29.0</td>
<td>9.7</td>
</tr>
</tbody>
</table>

[a] In °C. [b] In kcal/mol.
1.4 Origin of the Molecular Recognition Properties of Fluorinated Nucleobases

1.4.1 Stacking and desolvation: Insights from thermodynamic analyses

RNA stability is predominantly governed by base stacking, solvation forces and hydrogen bonding. The above data provide a rich source for analysing individual contributions of these forces, in particular, in view of the role of fluorine in molecular recognition.

The influence of base stacking was determined by comparing stabilities of duplex RNA with a respective base at position X, whereas the Y position is empty (abasic site) (Parsch and Engels, 2002). Compared to X = U, incorporation of the NNI 46 increases the duplex stability by 1.3°C (0.2 kcal/mol), demonstrating, as already found for ‘dangling end’ residues in the context of DNA (Guckian et al., 2000), that NNIs stack more strongly than their natural counterparts. Fluorine substitution then leads to another gain in stability through base stacking in that X = 1 increases the duplex stability by 2.7°C (0.5 kcal/mol) and X = 4 by 4.4°C (1.1 kcal/mol). Both an increase in the molecular dipole moment (46: 0.3 D, 1: 2.4 D, 4: 2.2 D) (Lai et al., 2003), resulting in increased van der Waals dispersive forces with the neighbouring bases (Lai et al., 2003) and a higher lipophilicity (46: log P = 1.05, 1: 1.50, 4: 1.70) account for this. Compared to X = G as a reference, X = 41 even leads to a stabilization by 5.4°C (1.2 kcal/mol), demonstrating that stacking interactions become more favourable with increasing size of the NNI, similar to the finding that a

<table>
<thead>
<tr>
<th>X</th>
<th>Y</th>
<th>$T_m^{[a]}$</th>
<th>$\Delta G^{[b]}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Benzene</td>
<td>Benzene</td>
<td>26.4</td>
<td>8.5</td>
</tr>
<tr>
<td>4FB</td>
<td>4FB</td>
<td>32.5</td>
<td>10.2</td>
</tr>
<tr>
<td>2,4DFB</td>
<td>2,4DFB</td>
<td>35.2</td>
<td>11.6</td>
</tr>
<tr>
<td>2,4,5TFB</td>
<td>2,4,5TFB</td>
<td>35.8</td>
<td>11.8</td>
</tr>
<tr>
<td>2,3,4,5TetFB</td>
<td>2,3,4,5TetFB</td>
<td>38.0</td>
<td>12.2</td>
</tr>
</tbody>
</table>

[a] In °C. [b] In kcal/mol.

Table 1.3. Thermodynamic stabilities of model 12mer RNA containing self-pairs of benzene and fluoro substituted benzenes (FB).
purine base stacks on a duplex more strongly than a smaller pyrimidine base (Petersheim and Turner, 1983). In the case of the fluoroindoles this could be determined by the dangling end method; for 25 a \( T_m \) increase of 11°C (2.7 kcal/mol) and for 27 of 4°C (0.9 kcal/mol) was measured (Bozilovic, 2008).

An \( X = U, Y = U \) mismatch (\( T_m = 30.1°C \)) was used as a reference for investigating the influence of desolvation of the hydrogen bond donors and acceptors of the natural bases during formation of the base pair with a NNI (Parsch and Engels, 2002). This desolvation effect was found to destabilize the RNA duplex in the case of \( X = 46, Y = U \) by −8.3°C (−2.2 kcal/mol). Notably, fluorine substitution reduced this effect to about −6.6°C (−1.7 kcal/mol) for both the \( X = 4 \) and 41 cases. A similar trend was found when base pairing between natural nucleobases and universal fluorinated NNIs was investigated by potential of mean force calculations (Koller et al., 2010), indicating interaction differences between paired bases, with more attractive interactions in the case of 4 and 41 than in the case of 46 (see below).

1.4.2 C–H...F–C interactions: Crystallographic analysis of fluoro substituted NNIs

When paired against natural bases, fluorinated NNIs destabilize DNA and RNA helices and exhibit universal base properties (Table 1.2) (Parsch and Engels, 2002; Lai and Kool, 2004). These observations make Watson–Crick base pairing involving hydrogen bonds to fluorine unlikely (Parsch and Engels, 2002; Somoza et al., 2006). When paired opposite one another, however, a considerable degree of stability is regained (Table 1.3) and a selective pairing of fluorinated NNIs in the context of nucleic acids is observed (Parsch and Engels, 2002; Lai and Kool, 2004; Kopitz et al., 2008). Weak C–F...H–C dipolar interactions have been implicated to act as stabilizing forces in this case (Lai and Kool, 2004; Kopitz et al., 2008). These findings relate to the present discussion as to whether ‘organic fluorine’ can act as a hydrogen bond acceptor and under which conditions (Dunitz and Taylor, 1997; Evans and Seddon, 1997; Guckian et al., 2000; Dunitz and Schweizer, 2006).

Following the idea that crystal packing information allows information to be deduced about non-bonded interactions (Velec et al., 2005),
initially crystal structures of fluorobenzene nucleosides 1–4 were determined. To our surprise the well-known herringbone pattern of benzene nucleosides was only present in nucleoside 2 (Matulicadamic et al., 1996). In contrast, for nucleosides 1, 3 and 4, when crystallized from water, the C–F···H–C distance between neighbouring aromatic rings of 1 is 238 pm, which is significantly shorter than the sum of the van der Waals radii of 255 pm (Rowland and Taylor, 1996). Similarly, when 1 was crystallized from methanol, a distance of 230 pm was found between the fluorine and a hydrogen atom in ortho position to F in a neighbouring aromatic ring (Parsch and Engels, 1999; Bats et al., 2000). The C–F···H–C arrangement shows a nearly linear configuration with a C–F···H angle of 158° (Fig. 1.8). We interpreted these findings in terms of weak attractive C–F···H–C dipolar interactions. The crystal structure of 3 revealed that the shortest C–F···H–C distance amounts to exactly the sum of the van der Waals radii of 255 pm of fluorine and hydrogen. This distance is found between the fluorine and H5′ of the sugar. Compared to the aromatic C–H bond adjacent to a fluorine-bound carbon in 1, the less polar aliphatic C–H5′ bond apparently provides less attraction for the aromatic fluorine.

Subsequently, we were able to determine the crystal structures of several fluorinated NNIs, including difluorobenzimidazoles and mono- and difluoroindoles (Parsch and Engels, 2002; Zivkovic and Engels, 2002; Bozilovic et al., 2007b). With respect to the latter, in 23 the distance

---

Figure 1.8. Crystal packing of 1′-deoxy-1′-(4-fluorophenyl)-β-D-ribofuranose 1.
between fluorine and a hydrogen in ortho position to F in a neighbouring molecule was larger than the sum of the van der Waals radii (269 pm). In turn, the shortest distance between those two atoms was found for 26 at 230 pm with a C–H···F angle of 124°. The crystal structures of 24, 25 and 27 showed intermediate distances of 246 pm, 240 pm and 239 pm, respectively (Bozilovic et al., 2007b).

Plotting the H···F distances and C–H···F angles according to Desiraju and coworkers (Thalladi et al., 1998) finally revealed that some of the C–H···F–C interactions found between fluorinated nucleosides have the shortest H···F distances reported so far for sp²-centred H···F interactions (Fig. 1.9).

![Figure 1.9](image_url)

**Figure 1.9.** F···H distances and C–F···H angles of C–F···H–C contacts found in crystal packings according to Thalladi et al. (1998) (black circles) as well as in crystal packings of fluorobenzene nucleosides (red crosses) (Parsch and Engels, 1999; Zivkovic and Engels, 2002), difluorobenzimidazole nucleosides (green crosses) (Zivkovic and Engels, 2002), and fluoroindole nucleosides (Bozilovic et al., 2007a).
1.4.3 Molecular dynamics simulations and free energy calculations

Additional insights at an atomic level into the stability determinants of RNA incorporating fluorinated nucleosides are provided by computer simulations. State-of-the-art molecular dynamics (MD) simulations and binding free energy calculations together with a structural component analysis were performed for RNA duplexes containing fluorobenzene self pairs at positions X and Y, resulting in deviations between experimental and computed relative binding free energies of less than 0.4 kcal/mol (Kopitz et al., 2008). Notably, these calculations revealed different origins for the incremental increase in duplex stability with increasing number of fluorine substituents in the NNI (Table 1.3): for the transitions \(46 \rightarrow 1\) and \(37 \rightarrow 38\) the binding free energy changes are dominated by favourable solvent contributions, resulting in an indirect effect of fluorine substitution. In contrast, for the transitions \(1 \rightarrow 4\) and \(4 \rightarrow 37\), changes of interactions within the RNA contribute favourably to the observed stability gain, showing a direct effect of fluorine substitution.

How can one explain these findings? Interestingly, global molecular properties such as the lipophilicity of the nucleosides or the molecular dipole moment of the NNI were inappropriate to explain the differences. Rather, the observed trend parallels differences in surface area regions contributed by fluorine atoms that are buried upon duplex formation (Fig. 1.10). This points to a local influence of fluorine substitution and can be explained by the poor aqueous solvation of C–F dipoles, yielding a hydrophobic character of these regions (Guerra and Bickelhaupt, 2003).

As for the direct RNA contributions to duplex stability, weak attractive C–F⋯H–C interactions between the self pairs were identified as stabilizing forces, with more short-range interactions present in 4 and 37 than in 1 (Fig. 1.11). This interpretation is corroborated by the analysis of the occupancy of C–F⋯H–C interactions: C–F⋯H–C interactions prevail for a larger fraction of time in the 4 (0.76) and 37 (0.70) cases than in the case of 1 (0.46).

Apparently, the fluorobenzene self pairs in the context of duplex RNA constitute a well-structured supramolecular system, which leads to favourable C–F⋯H–C interactions between self pairs of 4 and 37, as was found for other well-structured molecular environments such as enzyme active sites (Olsen et al., 2003; Olsen et al., 2004) or crystals (Desiraju, 2002; Reichenbacher et al., 2005).
Figure 1.10. Averaged structures obtained from MD trajectories of model 12mer RNA containing the self-paired fluorobenzene nucleotides 1 (a), 4 (b), 37 (c), and 38 (d). The solvent-accessible surfaces of the fluorine atoms are depicted as green meshes.

Figure 1.11. Averaged structures obtained from MD trajectories of model 12mer RNA containing the self-paired fluorobenzene nucleotides 1 (a) and 4 (b). (c) Frequency distributions of the distances between the H3 atom of the bases at position X and the F4 atoms of the bases at position Y for 1 (red), 2 (green), and 3 (blue). The distances computed for 1 and 4 are marked by a yellow dashed line in (a) and (b).
In another study, the ‘universal’ character of fluorinated NNI was investigated by potential of mean force calculations of base pairing between natural bases and fluorobases (A.N. Koller, J. Bozilovic, J.W. Engels, H. Gohlke, unpublished results). In agreement with previous studies (Stofer et al., 1999), Watson–Crick base pairing was computed to be favourable by about 1.5–2 kcal/mol per hydrogen bond formed. In contrast, pairing between 4 or 41 and the natural bases A or C was found to be unfavourable by 0.55–1.01 kcal/mol, in agreement with experiment (Parsch and Engels, 2002) and another MD study (Zacharias and Engels, 2004). Yet, for a given fluorinated NNI, the differences in the base pairing free energies with either one of the native bases are between 0.14 and 0.38 kcal/mol, supporting the universal base pairing properties. Finally, pairing between natural bases and 46 is more unfavourable by 0.6–1.0 kcal/mol than if a fluorinated base is used instead. Apparently, more attractive pairing interactions prevail in the case of 4 or 41 compared to 46. They arise from dipole–dipole interactions involving the C–F bond of the fluorobases and the exocyclic amine group of A or C and weak hydrogen bonds between N1 (N3) of A (C) and H–C3(5) of 4 (41).

### 1.5 Incorporation of Fluoro Substituted NNI into the Hammerhead Ribozyme and siRNA Constructs and their Acceptance by Polymerases

**1.5.1 Hammerhead ribozyme**

RNA as a central molecule in the chemistry of life is involved in the cellular process of gene expression and protein biosynthesis (Gesteland et al., 1999). RNA exhibits a great structural diversity and its secondary as well as tertiary structure is mainly stabilized by hydrogen bonding and base stacking. Many efforts have already been undertaken to elucidate structural changes via incorporation of artificial nucleosides.

Here, we intended to incorporate fluoro substituted NNI into a hammerhead ribozyme (Fig. 1.12) that is directed against the integrase region of the human immunodeficiency virus (HIV), a system we had introduced earlier (Klebba et al., 2000; Mueller-Kuller et al., in press).
Hammerhead ribozymes are catalytically active ribonucleic acids and interfere with gene expression through hydrolysis of the complementary mRNA. This makes them potential therapeutic agents for gene therapy (Lewin and Hauswirth, 2001; Scott, 2007). Recognition of viral mRNA and catalytic activity is dependent on Watson–Crick base pairing and decreases dramatically in the presence of point mutations (‘hot spots’) in the target region, since ribozymes are inactivated by mutating bases in the catalytic region. We rationalized that incorporation of universal bases in these positions should allow tolerance of escape mutants in HIV (Klöpffer and Engels, 2003, 2004, 2005).

The choice of the universal base was based on our investigations of fluorinated NNI in a model 12mer RNA, where the disubstituted fluorobenzene and -benzimidazole derivatives NNI 4 and 41 appeared to be ideal universal bases (Loakes, 2001). These NNIs did not differentiate thermodynamically between the four natural nucleosides A, C, G or U. Furthermore an enhanced base stacking ability upon additional fluorination was observed. This suggested choosing analogues 1 and 42 for the biological study. However, a destabilization of the modified RNA 12mer duplexes was noticed, which is due to the lack of hydrogen bonding interactions between the modified and natural bases. In an effort to compensate this reduced duplex stability, we synthesized the 2’-β-aminoethyl substituted fluorinated NNI 50 and 51 and investigated their ability to stabilize RNA duplexes. As the primary amino group is protonated under physiological
conditions, the intermolecular electrostatic interaction with the negatively charged RNA backbone of the second RNA strand as expected increased the stability of the RNA duplexes, probably due to charge–charge interactions (Klöpffer and Engels, 2004).

The hammerhead ribozyme was then modified with the universal bases 50 and 51 (Fig. 1.12). In order to investigate the ribozyme’s ability to tolerate point mutations in the target sequence without losing its catalytic activity, we analysed the kinetics of cleavage reactions. As expected, no discrimination with respect to the individual base pairing partner was found: although the overall efficiency was reduced by about one order of magnitude, resulting in a cleavage rate of 1 per 5 minutes, the rates towards A, U, C and G in the cognate mRNA were identical to within 30%. The difference between the fluorobenzene and fluorobenzimidazole ribosides was in favour of the benzimidazole, probably due to better stacking. Later on it was found that the hammerhead design can be improved using additional loop stabilization, but this was not pursued any further due to the identification of siRNA as potential gene regulators (Canny et al., 2004).

1.5.2 Fluorobenzene and benzimidazoles in RNA interference and siRNA

Since the recognition of RNA interference (RNAi) in 1998, the process by which specific mRNAs are targeted and degraded by complementary short-interfering RNAs (siRNA) became a powerful tool to control gene function (Fire, 2007; Mello, 2007). The generally accepted mechanism of RNAi can be divided into two main steps. In the first step, double-stranded RNA (dsRNA) is cleaved into short 21–24nt siRNAs (Elbashir et al., 2001). This process is catalysed by Dicer, an endonuclease of the RNase III family. The resultant siRNA duplexes have 3’-overhangs of 2nt with 3’-hydroxyl termini and a 5’-phosphate at both ends. In the second step, siRNAs are incorporated into the RNA-induced silencing complex (RISC). A helicase in RISC unwinds the duplex siRNA, which then pairs to messenger RNAs (mRNAs) that bear a high degree of sequence complementarity to the siRNA. In humans the degradation of the target mRNA is mediated by the Argonaute 2 protein associated with RISC. The target mRNA is cleaved in the complementary region at the phosphodiester bond that lies across
from nucleotides 10 and 11 of the 5′-end of the siRNA (Fig. 1.13). For RNAi-mediated mRNA cleavage and degradation to be successful, a 5′-phosphate must be present on the antisense strand and the double helical antisense-target mRNA duplex must be in the A-form. The X-ray structure of Ago2 and either a single or double stranded oligonucleotide highlights the structural situation (Wang et al., 2008).

Chemically modified nucleosides have been shown to be of great importance for antisense strategies and are now being applied for RNA interference-mediated gene silencing (Bramsen et al., 2009). Since the incorporation of 2′-amino modified NNIs had been shown to yield active ribozymes, the analogous constructs were designed for siRNAs. During our studies several publications appeared showing the use of fluorobenzene ribosides in this case. First, Kool and coworkers used compound 4 in a full seed walk (Fig. 1.13) of siRNA against Renilla luciferase (Somoza et al., 2006). From these data it is obvious that the 3′- and 5′-ends tolerate the fluorobenzene well. The central position 10 is highly discriminative. Surprisingly, position 7 gave a high activity too, indicating a possible interaction with the RISC complex (Fig. 1.13). A similar study by the Alnylam group was in good agreement, even though they incorporated the ribo-difluorotoluyl nucleotide (Xia et al., 2006). In our hands, the central positions from 9 to 11 are also less tolerant for fluorobenzene 4 and other tested fluorobenzimidazole nucleosides. When incorporating fluorobenzimidazole 41 at position 21 of siRNA the activity was still preserved, which indicates a successful phosphorylation at the 5′-end. This proved to be particularly interesting because these fluorobenzimidazole ribosides can be easily derivatized by a sequence of Michael addition with acrylonitrile followed by Raney-nickel reduction yielding the aminopropyl derivative (Haas and Engels, 2007).
1.5.3 Polymerase acceptance of fluoro benzimidazoles

In collaboration with R. Kuchta at Boulder, Colorado, we tested our fluoro- 
rinated benzimidazole nucleoside analogue of 41 (deoxynucleoside) 
against polymerases. These were DNA polymerase α (pol α) and Klenow 
fragment (exo-) of DNA polymerase I (Escherichia coli). Both pol α and 
Klenow fragment exhibit a remarkable inability to discriminate against 
these analogues as compared to their ability to discriminate against incor-
rect natural deoxynucleotide triphosphates (dNTPs). Neither polymerase 
shows any distinct electronic or steric preferences for analogue incorpor-
ation (Kincaid et al., 2005).

Another set of analogues was designed to examine human DNA prim-
ase, which synthesizes short RNA primers that DNA polymerase α 
further elongates. Primase readily misincorporates the natural nucleotide 
triphosphates (NTPs), which generates a wide variety of mismatches. In 
contrast, primase exhibited a remarkable resistance to polymerizing NTPs 
containing NNI. This is different from other polymerases where the shape 
concept, put forward by Kool, is more applicable (Kool and Sintim, 2006). 
We tested bases whose shape was almost identical to the natural bases 
(4-aminobenzimidazole and 4,6-difluorobenzimidazole) (Klöpffer and 
Engels, 2005), bases with very different shapes compared to natural bases 
(5- and 6-trifluoromethylbenzimidazoles 43 and 44), bases much more 
hydrophobic than natural bases (4- and 7-trifluoromethylbenzimidazole 
42 and 45), bases with hydrophobicities similar to natural bases but with 
the Watson–Crick hydrogen bonding groups in unusual positions (7-β-D-
guanine) and bases capable of forming only one Watson–Crick hydrogen 
bond with the template base (purine and 4-aminobenzimidazole). 
Primase was found only to polymerize NTP analogues capable of forming 
Watson–Crick hydrogen bonds, which explains the failure for the incor-
poration of 42–45.

1.6 Conclusion

We have described the syntheses of fluorinated benzene, benzimidazole 
and indole nucleobase analogues and their incorporation into model 
RNA, ribozymes and siRNA. The analogues act as NNIs and allow probing
of the physical forces that govern the stability of RNA. Notably, several of these NNIs showed a lack of discrimination against natural bases and thus behave as universal bases. Furthermore, the stability of model RNA incorporating self-pairs of fluorinated NNIs increased with the number of fluorinated substituents and reached that of a natural base pair. Combined crystallographic studies, thermodynamic analyses and computer simulations furthermore shed light on the role of organic fluorine in molecular recognition. These studies demonstrated an intricate influence of the molecular environment in this case. As a consequence, it may generally not be sufficient to discuss the molecular recognition properties of organic fluorine in terms of global molecular descriptors. Rather, analyses at an atomic level are required.

Acknowledgements

We are grateful for the tremendous efforts of our coworkers on this project. This work was funded by the Deutsche Forschungsgemeinschaft within the Sonderforschungsbereich 579 (‘RNA–ligand interactions’) and by a grant to J.W.E. within the EU-FP6 RIGHT project (LSHB-CT-2004-005276). H.G. gratefully acknowledges support by the Frankfurt Centre for Scientific Computing.

References


