DNA Structures

Fluorinated DNA Bases as Probes of Electrostatic Effects in DNA Base Stacking**

*Jacob S. Lai, Jin Qu, and Eric T. Kool**

The noncovalent interactions affecting the thermodynamic stability of natural and modified DNA have been topics of broad interest in recent years. The effects of sterics, stacking, hydrogen bonding, and minor-groove solvation have been considered as contributing factors.[1–4] Probably the dominant stabilizing factor in helical DNA is base stacking.[5,6] In order to probe the physical factors that contribute to the stability of this stacking in water, measured melting data of short DNA oligomers, both naturally and nonnaturally substituted, has been studied.[5,7,8] Such experiments have suggested that van der Waals and solvophobic forces can be important contributors to the stabilization of stacking. Beyond this, theoretical work has pointed out the possible importance of electrostatic interactions in the stability and preferred geometry of stacked bases in DNA.[9–14] Understanding these issues could allow for better design of modified DNAs, but relatively little experimental information is available on such electrostatic factors.

[*] Prof. Dr. E. T. Kool, J. S. Lai, Dr. J. Qu
Department of Chemistry, Stanford University
Stanford, CA 94305-5080 (USA)
Fax: (+1) 650-725-0259
E-mail: kool@leland.stanford.edu

[**] This work was supported in part by the U.S. National Institutes of Health (GM52956/EB002059).

Supporting information for this article is available on the WWW under http://www.angewandte.org or from the author.

It is well documented that aromatic rings can exhibit significant dipolar and quadrupolar electrostatic interactions in certain environments. Studies in nonpolar solvents using aromatic compounds containing various electron-donating and -withdrawing substituents have demonstrated significant electrostatic effects in stacking energetics and geometries.[15–17] Although in aqueous systems the electrostatic dipole effects are greatly diminished, localized electrostatic effects are still believed to play a role in governing neighboring base-pair geometries in DNA.[6,12] Studies with small-molecule model systems in water have suggested that electrostatic effects are relatively weak, and that dispersive effects are a major factor governing stacking stability.[18,19] Hydrophobic effects in such model systems appear to play only a small role, although this issue has been debated.[18,20,21]

Quadrupolar interactions have also been documented in specialized cases; for example, benzene is capable of electrostatic interactions with molecules containing a positive charge, as demonstrated by well-documented cation–π interactions, even in aqueous systems.[22] By contrast, perfluorobenzene, with its opposite quadrupolar sign, can stack well (in low polarity environments) with electron-rich aromatic rings.[23] A recent computational study also showed the coordination by water in interactions with benzene and perfluorobenzene.[24] Although benzene-,[25] 2,4-difluorobenzene-,[26] 2,4,5-trifluorobenzene-,[26] and pentafluorobenzene-substituted[27] deoxyribonucleosides and 4-monofluorobenzene ribonucleoside[28] have been previously described, no information on their relative stacking abilities is available, nor is there any data on their interactions with varied neighboring DNA bases.

We now describe a series of fluorinated aromatic nucleoside analogues having a wide range of dipole and quadrupole moments. We have studied the stacking thermodynamics of these compounds in short synthetic DNA duplexes, with all four neighboring nucleobases. The results shed light on the importance and origins of electrostatic interactions in DNA base stacking, and reveal some previously unrecognized structural and electrostatic effects that will be useful in future molecular designs.

The seven deoxynucleosides studied here are shown in Figure 1. In all cases, the deoxyribose moiety is constant, but the "base" groups vary in the extent and orientation of fluorine substitution, from zero substitutions (benzene) to the maximum of five (pentafluorobenzene). These compounds were prepared by treating the appropriate lithiated aromatic species with an O-protected deoxyribonolactone derivative.[29]

Figure 1b depicts calculated electrostatic surface potentials of the aromatic base analogues, in which the effects of the deoxyribose 1′-carbon are approximated with an attached methyl group. The electrostatics vary widely over the series, with the phenyl nucleoside showing a negative potential at the center of the flat aromatic face, and the pentafluorinated case having a strongly positive potential. Thus the quadrupoles are gradually inverted over this series. By comparison, natural DNA bases are not as strongly polarized (in the quadrupolar sense), and the potentials are generally close to neutral (similar to monofluorobenzene, see the Supporting Informa-

 DOI: 10.1002/ange.200352531

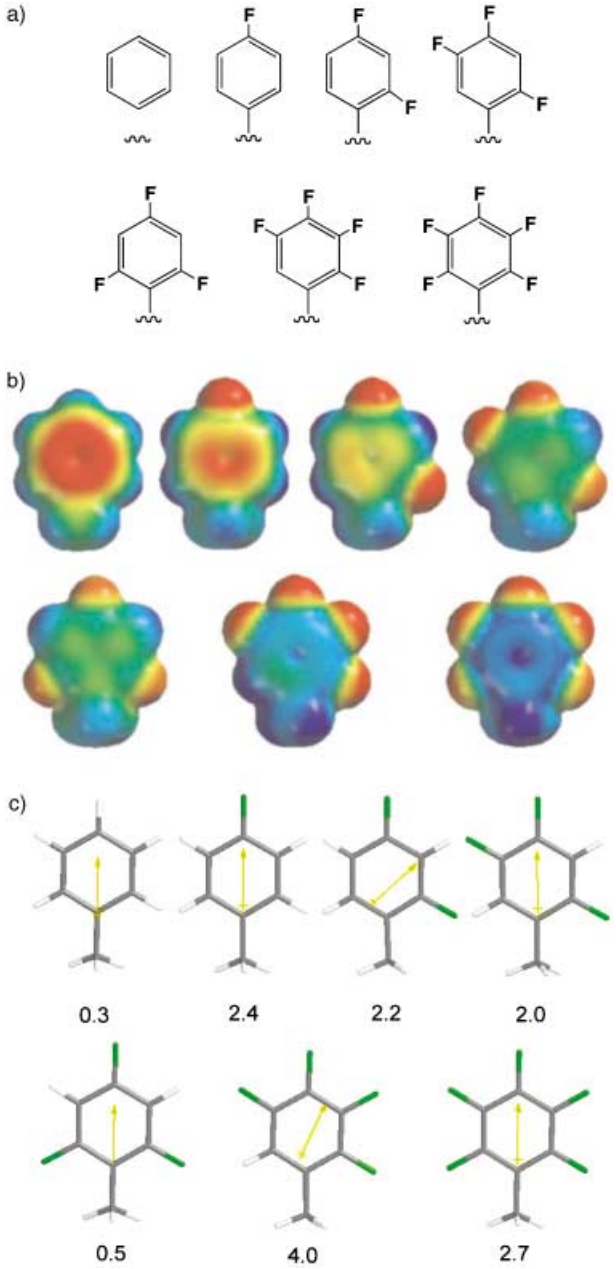

Figure 1. Structures and properties of the seven nucleoside analogues in this study. a) Molecular structures of the nucleobases (all are attached at C1′ of the deoxyribose). b) Calculated electrostatic surface potentials of six progressively fluorinated aromatic base analogues, with an attached methyl group to approximate the effects of the deoxyribose 1′-carbon (red depicts negative potential and blue, positive). c) Calculated dipole moments (debye) of aromatic fluorinated base analogues; dipole orientations are shown with yellow arrows, fluorine atoms are in green. Electrostatics were calculated with Spartan'02 (Wavefunction Inc.) employing the AM1 Hamiltonian.

tion). In addition to differences in quadrupoles, this series has a broad range of dipole moments (Figure 1c), ranging from a calculated 0.4 (for the phenyl case) to 4.5 debye (for the tetrafluorinated compound). Generally, the dipole directional orientations are quite similar over the series (Figure 1c). By comparison, natural DNA bases have dipole moments that are generally large, but with orientations that vary widely (Supporting Information). Finally, to test the effects of location of the fluorine substituents on the base, we studied two different trifluoro-substituted analogues: the 2,4,5-trifluorophenyl and 2,4,6-trifluorophenyl deoxyribonucleosides.

The seven nucleoside analogues were incorporated into synthetic oligonucleotides to study their stacking propensities with natural DNA bases as neighbors. They were incorporated by standard methods on an automated DNA synthesizer, and were characterized in DNA by NMR spectroscopy and by mass spectrometry.

The dangling-end experimental method[30,31] was utilized to evaluate the ability of the fluorinated aromatic base analogues to stabilize DNA duplexes when placed directly adjacent to the helix. The relative stabilizations were measured by comparison of the energetics for helix–coil melting transition of the "core" duplex (lacking any "dangling" nucleotide) to that containing the extra nucleotide. Five short self-complementary sequence contexts were used; the sequences were chosen because they had been shown previously to be well-behaved thermodynamically in the dangling-end configuration.[7,32] Data from two of the contexts are shown in Table 1, and data for the remaining three (showing similar trends) are given in the Supporting Information. By use of these five contexts we were able to compare stacking effects of the base analogues with all four neighboring DNA bases in the helix.

Thermodynamics were obtained both by curve fitting and by the van't Hoff method. All the duplexes appeared to behave in a two-state fashion and had well-shaped melting curves indicative of cooperative interactions of the dangling ends (data not shown). All of the seven unnatural bases displayed significant stabilization of the duplexes relative to the core sequences (Table 1). The least stabilized is the case with a dangling pentafluorophenyl nucleotide on the $(dCGCGCG)_2$ core duplex, which gives an increase in T_m of only 2.9 °C and contributes $-0.5\ kcal\,mol^{-1}$ of stability ("$\Delta\Delta G°$ stacking" in Table 1). Almost as poorly stabilizing is the 2,4,6-trifluorinated case, which will be discussed below. The largest stabilizing interaction is observed with the 2,3,4,5-tetrafluorophenyl dangling nucleotide in that same sequence, with a T_m increase of 12.6 °C and a large stabilization of $-2.2\ kcal\,mol^{-1}$ compared to the unsubstituted core sequence.

Figure 2 shows the T_m data graphically, illustrating trends over the series with lines connecting data points. The "C adjacent" and "A adjacent" series have the most data available; here we see that mono-, di-, and trifluorinated bases (omitting, for the moment, the 2,4,6-trifluoro isomer) all stabilize the core sequences with similar propensity. For most cases the previously unknown tetrafluoro analogue stacks somewhat more stably than the rest, while in all cases the previously known pentafluoro compound is surprisingly poor at stacking, behaving quite differently than the other compounds across the series.

It is worth noting that simple placement in a 5′-dangling position does not always guarantee a stacked orientation with the neighboring DNA bases. However, we expect that, since these analogues are relatively similar in size and geometry, their propensities for preferred geometries in DNA might

Table 1: Stacking of fluoroaromatic nucleotides as measured by thermal denaturation studies in two sequence contexts.[a]

Dangling residue	T_m [°C][b]	ΔT_m [°C]	$-\Delta H$ [kcal][c]	$-\Delta S$ [eu][c]	$-\Delta G_{37}$ [kcal][c]	$-\Delta G_{37}$ [kcal][d]	$\Delta\Delta G°$ stacking
XCGCGCG							
none (core duplex)	41.4	—	43.6	115	8.0±0.1	8.1±0.1	
phenyl	49.1	7.8	44.7	115	9.2±0.1	9.4±0.1	1.2±0.1
4-fluorophenyl	51.2	9.8	64.9	175	10.3±0.3	9.9±0.2	2.0±0.2
2,4-difluorophenyl	52.5	11.2	51.7	134	10.0±0.1	10.2±0.1	
2,4,5-trifluorophenyl	52.2	10.8	55.0	145	10.1±0.3	10.0±0.1	2.0±0.2
2,4,6-trifluorophenyl	45.3	4.0	37.3	104	8.4±0.2	8.6±0.1	0.4±0.1
2,3,4,5-tetrafluorophenyl	53.9	12.6	53.8	93	10.3±0.2	10.2±0.2	2.2±0.2
pentafluorophenyl	44.2	2.9	39.1	99	8.5±0.2	8.7±0.1	0.5±0.2
XACAGCTGT							
none (core duplex)	40.0	—	44.8	119	7.9±0.1	8.1±0.1	
phenyl	44.4	4.4	57.6	157	8.9±0.2	9.0±0.1	1.0±0.1
4-fluorophenyl	48.7	8.7	66.6	183	9.9±0.1	9.6±0.1	1.8±0.1
2,4-difluorophenyl	47.8	7.8	62.4	170	9.7±0.2	9.6±0.1	1.7±0.1
2,4,5-trifluorophenyl	48.9	9.0	67.8	186	10.1±0.2	9.7±0.1	1.9±0.2
2,4,6-trifluorophenyl	47.8	7.9	53.6	143	9.3±0.2	9.4±0.2	1.3±0.2
2,3,4,5-tetrafluorophenyl	50.9	10.0	76.3	212	10.7±0.2	9.8±0.2	2.3±0.2
pentafluorophenyl	47.3	7.0	55.2	148	9.3±0.0	9.3±0.2	1.3±0.1

[a] Free energy of stacking ($\Delta\Delta G°$) is calculated as the difference between the free energies of the duplexes containing dangling residues from the energy of the core duplex. Data from three other sequence contexts is given in the Supporting Information. [b] Conditions: 1 M NaCl, 10 mM sodium phosphate pH 7.0; 5.0 μM DNA-strand concentration for the T_m value shown. [c] Thermodynamic values calculated from van't Hoff plots. [d] Average free energy from fits to individual melting curves.

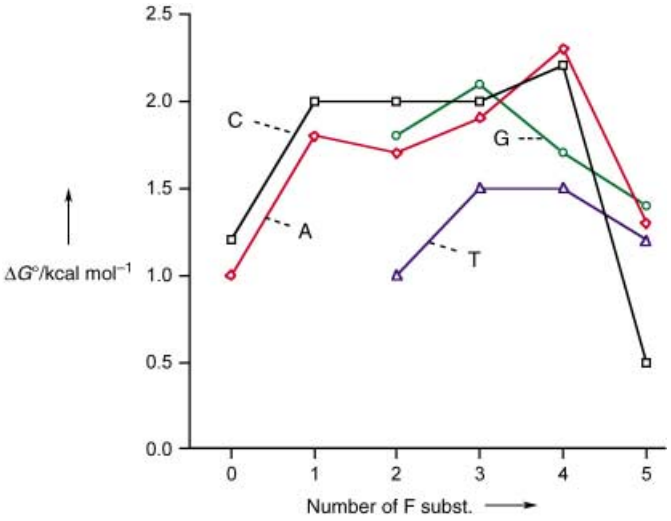

Figure 2. Plot showing trends in stacking free energies as a function of the number of fluorine substitutions on the phenyl deoxyribose adjacent to cytosine (black), adenine (red), guanine (green), and thymine (blue). Primary data are given in Table 1 and in Table S1 in the Supporting Information; the 2,4,6-trifluorinated case is omitted here.

also be similar. It will be important in the future to confirm geometries by structural studies, particularly in a strongly stabilizing case such as the tetrafluorinated analogue, as well as in a poorly stabilized (and potentially distorted) case such as with pentafluorobenzene. Regardless of geometry, however, the present results do indicate which structures are most stabilizing for future molecular designs.

We hypothesized that steric factors might contribute to the poor stabilization by the pentafluorinated species. Although fluorine is a relatively small substituent, it is generally accepted that even small groups can alter the glycosidic orientational preference in nucleosides, presumably by steric interactions with neighboring bonds and substituents.[33,34] To test this further we prepared a second trifluorinated species, this one with 2,4,6 substitution, for direct comparison to the 2,4,5-substituted case. Measurement in the dangling-end contexts revealed that, like the pentafluorinated case, the bis-*ortho*-substituted trifluorobenzene case was very poor at stabilization (Table 1). This is a remarkable difference: a change of 6.8 °C (1.6 kcal) on moving one fluorine atom from the *ortho* to the *meta* position.

Thus the data establish that bis-*ortho* substitution by even small fluorine groups can have a surprisingly large effect on stabilization, causing stacking of both the pentafluoro- and 2,4,6-trifluoronucleosides to be disrupted. We hypothesize that this may be due to a sterically induced twist in the glycosidic bond and/or in the sugar; structural studies will be helpful in evaluating this in the future. This finding explains the previously observed strong destabilizations seen in DNAs containing pentafluorobenzene.[27]

There is a gradual inversion of electrostatic potential at the centers of the flat aromatic surfaces going from the phenyl nucleoside to the pentafluorophenyl nucleoside across this series (see the electrostatic potential maps in Figure 1b). It was anticipated that this difference in electrostatic potential might be a significant factor in the stacking of these compounds in duplex DNA, particularly with differing natural adjacent bases. Electrostatic maps of the four natural DNA bases suggest that adenine is the most electron-rich, while thymine is the least (see the Supporting Information), although the differences are relatively small. However, from the T_m and $\Delta G°$ differences observed here (Table 1 and

Figure 2)), there appears to be relatively poor correlation between numbers of fluorine groups and stabilization. For example, in the "C adjacent" and "A adjacent" series, the mono-, di-, and trisubstituted cases show nearly the same stacking propensities (Figure 3a). (Note that we omit the two bis-*ortho* cases from the analysis because of their unusual steric effects.) Since the tetrafluorinated species does stack somewhat more strongly, we cannot entirely rule out a quadrupolar effect, but the data suggest (see below) that simple dipole effects may provide a more consistent explanation for the results. Thus we conclude that, when at least one natural DNA base is involved, quadrupole effects are small or nonexistent, even with one strongly polarized partner. It remains to be seen, however, whether *two* adjacent non-natural stacking partners (which could be more strongly polarized than natural bases are) might exhibit quadrupolar stabilization or destabilization in water.

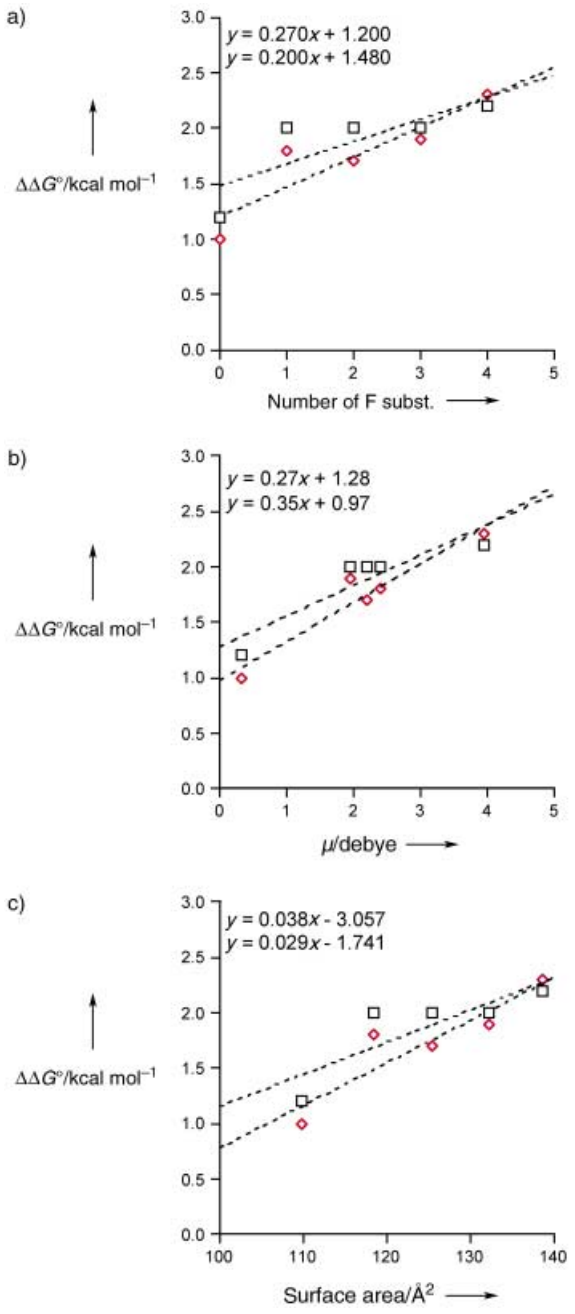

Figure 3. Testing possible linear relationships between stacking free energies and calculated physical properties of aromatic analogues in two comparative sequences with C adjacent (black) and with A adjacent (red) (see Table 1). a) Number of fluorine substitutions (a rough measure of quadrupole strength). b) Dipole moment μ (debye) of the methylated base. c) Estimated surface area of the dangling residue excluded from solvent on stacking. The two compounds with bis-*ortho* effects were omitted.

Dipole moments of the aromatic nucleobase analogues were calculated and plotted against ΔT_m and $\Delta\Delta G°$ (stacking) for both the "C adjacent" and "A adjacent" series (Figure 3b). Overall, it does appear that there is a linear correlation between dipole moments and stacking stabilization across these series. However, if permanent dipole effects are real, then the orientation of the analogue dipoles relative to the dipole orientations of adjacent bases should play a role in the electrostatic effects. Since the analogue dipole orientations are all similar (Figure 1c), one should compare the dipole orientations of the neighboring bases, which differ more greatly. The dipole directions for adenine and cytosine are oriented roughly 45° relative to one another (see the Supporting Information). If one assumes that a 5′-stacked base takes on the standard B-form conformation, then the dipole orientations of the analogues should be nearly opposed (180°) to that of a cytosine in the neighboring helix, whereas they should be only partially opposed to that of a neighboring adenine. Also possibly significant is the much stronger dipole of C (6.0 debye as the 1-methyl derivative) relative to that of A (2.3 debye). These factors lead to the prediction of stronger dipole effects for a C neighbor than for an A neighbor. However, the plot (Figure 3b) shows very similar slopes for the two series, which is not consistent with these predictions.

A dipole in an end-stacked nucleobase can have two types of electrostatic effects on stacking: as a direct electrostatic interaction with a nearby permanent dipole, and also in the dispersive sense, by inducing an opposing dipole in the neighboring base. One would predict a stronger dispersive effect for a dangling base with a neighboring A than with a neighboring C, because of the greater polarizability of A. However, in the converse sense, C should induce a stronger dipole in the dangling base because of C's strong dipole. These opposing effects might tend to make C and A somewhat similar in stacking abilities, which is consistent with the similar slopes in Figure 3b.

Thus we tentatively conclude that, while dipole effects appear to be significantly stabilizing to DNA base stacking, the origin of the effect may lie in their contribution to van der Waals dispersive forces. Overall, dipole effects can explain only roughly half (and likely less) of the stabilization observed here. Extrapolation to zero dipole still leaves about half of the stacking stabilization intact. Moreover, the compound with largest dipole (the tetrafluorobenzene case) also has greater surface area than the parent benzene compound, a factor that likely also contributes favorably. Plots of surface area vs. stacking (Figure 3c) do show an apparent loose correlation of stabilization with surface area. A correlation of stacking with surface area has been reported

previously.[5] It remains to be seen whether the favorable effects of size in this case are due to solvophobic or dispersive effects.

Overall, our data are consistent with the notion that dispersive van der Waals attractions may be among the most important factors in DNA base stacking. The current results suggest that the electrostatic effects of nucleobase dipoles are significant in stabilizing stacking, but may explain only one-third to one-half of the stabilization for bases with strong dipoles. We further suggest that this dipole effect may be a result of dispersive induced-dipole attractions rather than of attractions between permanent dipoles. This leads to the suggestion that aromatic bases with large size and large dipole may be generally well-suited for stacking. Overall the data suggest that quadrupole interactions appear to be small for the natural bases, which have weak quadrupole moments. Finally, a previously unrecognized bis-*ortho* difluoro substitution effect was the largest factor observed in this series. This effect is clearly to be avoided in future designs of base analogues for helix stabilization.

Experimental Section

All synthetic methods and characterizations of compounds, oligonucleotide synthesis and characterizations, and thermal-denaturation methods and data are reported in the Supporting Information.

Received: July 31, 2003 [Z52531]

Keywords: DNA structures · electrostatic interactions · fluorine · pi interactions

[1] S. Smirnov, T. J. Matray, E. T. Kool, C. de los Santos, *Nucleic Acids Res.* **2002**, *30*, 5561 – 5569.
[2] M. L. Waters, *Curr. Opin. Chem. Biol.* **2002**, *6*, 736 – 741.
[3] a) J. Sponer, J. Leszczynski, P. Hobza, *J. Mol. Struct. THEOCHEM* **2001**, *573*, 43 – 53;b) J. Sponer, J. Leszczynski, P. Hobza, *Biopolymers* **2001**, *61*, 3 – 31.
[4] U. Rychlewska, B. Warzajtis, *Acta Crystallogr. Sect. B* **2001**, *57*, 415 – 427.
[5] K. M. Guckian, B. A. Schweitzer, R. X. F. Ren, C. J. Sheils, D. C. Tahmassebi, E. T. Kool, *J. Am. Chem. Soc.* **2000**, *122*, 2213 – 2222.
[6] C. A. Hunter, *J. Mol. Biol.* **1993**, *230*, 1025 – 1054.
[7] a) S. Bommarito, N. Peyret, J. SantaLucia, Jr., *Nucleic Acid Res.* **2000**, *28*, 1929 – 1934; b) J. SantaLucia, Jr., H. T. Allawi, P. A. Seneviratne, *Biochemistry* **1996**, *35*, 3555 – 3562.
[8] S. Nakano, Y. Uotani, S. Nakashima, Y. Anno, M. Fujii, N. Sugimoto, *J. Am. Chem. Soc.* **2003**, *125*, 8086 – 8087.
[9] V. K. Misra, B. Honig, *Biochemistry* **1996**, *35*, 1115 – 1124.
[10] S. C. Harvey, C. L. Wang, S. Teletchea, R. Lavery, *J. Comput. Chem.* **2003**, *24*, 1 – 9.
[11] C. A. Hunter, X. J. Lu, *J. Mol. Biol.* **1997**, *265*, 603 – 619.
[12] P. Hobza, J. Sponer, *J. Am. Chem. Soc.* **2002**, *124*, 11 802 – 11 808.
[13] a) J. Sponer, I. Berger, N. Spackova, J. Leszczynski, P. Hobza, *J. Biomol. Struct. Dyn.* **2000**, *2*, 383 – 407; b) J. Sponer, H. A. Gabb, J. Leszczynski, P. Hobza, *Biophys. J.* **1997**, *73*, 76 – 87.
[14] R. Luo, H. S. R. Gilson, M. J. Potter, M. K. Gilson, *Biophys. J.* **2001**, *80*, 140 – 148.
[15] a) F. Cozzi, F. Ponzini, R. Annunziata, M. Cinquini, J. S. Siegel, *Angew. Chem.* **1995**, *107*, 1092 – 1093; *Angew. Chem. Int. Ed. Engl.* **1995**, *34*, 1019 – 1020; b) F. Cozzi, R. Annuziata, M. Cinquini, J. S. Siegel, *J. Am. Chem. Soc.* **1993**, *115*, 5330 – 5331; c) F. Cozzi, M. Cinquini, R. Annunziata, T. Dwyer, J. S. Siegel, *J. Am. Chem. Soc.* **1992**, *114*, 5729 – 5733.
[16] F. J. Carver, C. A. Hunter, D. J. Livingstone, J. F. McCabe, E. M. Seward, *Chem. Eur. J.* **2002**, *8*, 2848 – 2859.
[17] M. J. Rashkin, M. L. Waters, *J. Am. Chem. Soc.* **2002**, *124*, 1860 – 1861.
[18] T. J. Liu, H. J. Schneider, *Angew. Chem.* **2002**, *114*, 1418 – 1420; *Angew. Chem. Int. Ed.* **2002**, *41*, 1368 – 1370.
[19] B. S. Palm, I. Piantanida, M. Zinic, H. J. Schneider, *J. Chem. Soc. Perkin Trans. 2* **2000**, *2*, 385 – 392.
[20] L. F. Newcomb, S. H. Gellman, *J. Am. Chem. Soc.* **1994**, *116*, 4993 – 4994.
[21] R. A. Friedman, B. Honig, *Biophys. J.* **1996**, *71*, 3525 – 3526.
[22] a) J. C. Ma, D. A. Dougherty, *Chem. Rev.* **1997**, *97*, 1303 – 1324; b) D. A. Dougherty, *Science* **1996**, *271*, 163 – 168.
[23] E. A. Meyer, R. K. Castellano, F. Diederich, *Angew. Chem.* **2003**, *115*, 1244 – 1287; *Angew. Chem. Int. Ed.* **2003**, *42*, 1210 – 1250.
[24] M. Raimondi, G. Calderoni, A. Famulari, L. Raimondi, F. Cozzi, *J. Phys. Chem. A* **2003**, *107*, 772 – 774.
[25] N. C. Chaudhuri, R. X. F. Ren, E. T. Kool, *Synlett* **1997**, 341 – 347.
[26] Z. X. Wang, W. Duan, L. I. Wiebe, J. Balzarini, E. De Clercq, E. E. Knaus, *Nucleosides Nucleotides Nucleic Acids* **2001**, *20*, 11 – 40.
[27] G. Mathis, R. Hunziker, *Angew. Chem.* **2002**, *114*, 3335 – 3338; *Angew. Chem. Int. Ed.* **2002**, *41*, 3203 – 3205.
[28] J. Parsch, J. W. Engels, *Helv. Chim. Acta* **2000**, *83*, 1791.
[29] U. Wichai, S. A. Woski, *Org. Lett.* **1999**, *1*, 1173 – 1175.
[30] M. Petersheim, D. H. Turner, *Biochemistry* **1983**, *22*, 256 – 263.
[31] D. H. Turner, N. Sugimoto, S. M. Freier, *Annu. Rev. Biophys. Biophys. Chem.* **1988**, *17*, 167 – 192.
[32] K. M. Guckian, B. A. Schweitzer, R. X. F. Ren, C. J. Sheils, P. L. Paris, D. C. Tahmassebi, E. T. Kool, *J. Am. Chem. Soc.* **1996**, *118*, 8182 – 8183.
[33] M. S. Cooke, K. E. Herbert, P. C. Butler, J. Lunec, *Free Radical Res.* **1998**, *28*, 456 – 469.
[34] J. Cadet, R. Ducolomb, C. Taieb, *Tetrahedron Lett.* **1975**, 3455 – 3458.

