Solution Structure of a Nonpolar, Non-Hydrogen-Bonded Base Pair Surrogate in DNA

ARTICLE *in* JOURNAL OF THE AMERICAN CHEMICAL SOCIETY · JULY 2000

Impact Factor: 12.11 · DOI: 10.1021/ja994164v · Source: PubMed

3 AUTHORS, INCLUDING:

Kevin M Guckian
Biogen Idec
20 PUBLICATIONS **1,685** CITATIONS
SEE PROFILE

Eric T Kool
Stanford University
280 PUBLICATIONS **13,206** CITATIONS
SEE PROFILE

Published in final edited form as:
J Am Chem Soc. 2000 July 26; 122(29): 6841–6847. doi:10.1021/ja994164v.

Solution Structure of a Nonpolar, Non-Hydrogen-Bonded Base Pair Surrogate in DNA

Kevin M. Guckian, **Thomas R. Krugh**[*], and **Eric T. Kool**[*,†]
Contribution from the Department of Chemistry, University of Rochester, Rochester, New York 14627

Abstract

We describe the structure in aqueous solution of a DNA duplex containing a base pair that is structurally analogous to A–T but which lacks hydrogen bonds. Base analogues F (a nonpolar isostere of thymine) and Z (a nonpolar isostere of adenine) are paired opposite one another in a 12 base pair duplex. The sequence context is the binding site of recently studied transcription factor hSRY. The Z–F pair has been shown to be replicated surprisingly well and selectively by DNA polymerase enzymes, considering that it is destabilizing and lacks Watson–Crick hydrogen bonds. The enzymatic studies led to the suggestion that part of the functional activity arises because the pair resembles a natural one in geometry. The present results show that, despite the absence of Watson–Crick hydrogen bonds, the Z–F pair structurally resembles an A–T pair in the same context. This lends support to the proposal that shape matching is an important component in replication, and suggests the general utility of using Z–F as a nonpolar replacement for A–T in probing protein–DNA interactions.

Introduction

Recent studies of deoxynucleosides having nonpolar base replacements have demonstrated that Watson–Crick hydrogen bonds are not required for a DNA polymerase to replicate a base pair. For example, a 2,4-difluorotoluene deoxynucleoside (abbreviated F), which acts as a nonpolar shape mimic for thymidine,[1] was found to act as a highly specific substrate for multiple DNA polymerases.2⁻4 The Klenow fragment of *Escherichia coli* DNA polymerase I was found to synthesize pairs of F with adenine with efficiency approaching that of a natural T–A base pair.[3,4] This was surprising because the F "base" shows little or no tendency to form a hydrogen-bonded pair with adenine,[5] and indeed, is quite destabilizing when paired with adenine in a DNA duplex.[3] More recently, experiments in other laboratories have also demonstrated efficient replication of non-H-bonded pairs.[6] Since hydrogen bonds are apparently dispensable in maintaining natural enzyme activity, we proposed that the fidelity of DNA replication might arise in large part from steric matching of bases within a tightly confined active site.[3,7,8]

This hypothesis was criticized by suggesting that the reason that F behaved as such an effective mimic of T was that it actually did form robust hydrogen bonds with adenine.9 After this, several experiments and calculations with this fluorinated analogue and related compounds have concluded that C–F…H–N hydrogen bonding character is quite poor in this context, especially in the aqueous environment.10⁻18 Nonetheless, a clearer way to avoid the issue of possible hydrogen bonds between F and A is to replace adenine with a nonpolar isostere, as

[*] To whom correspondence should be addressed. T. R. Krugh, krugh@chemistry.rochester.edu. E. T. Kool, kool@stanford.edu.
[†]Current address: Department of Chemistry, Stanford University, Stanford, California 94305-5080.

was done for thymine. To this end, we chose a 4-methylbenzimidazole replacement used for adenine (Z) and studied its ability to be replicated opposite F.[19,20] It was found, interestingly, that this pair was an efficient substrate for the Klenow polymerase, and it was processed with significant levels of selectivity.[21] On the basis of these results we suggested that steric accommodation within the DNA double helix is necessary and sufficient for efficient replication of a pair, and that it can also account for some (although not necessarily all) of the selectivity observed in replication as well.

A steric matching hypothesis for DNA replication[7] requires that to be a good substrate, a base pair (whether it consists of natural or nonnatural bases) must be able to adopt a geometry close to that of the standard pairs without steric clashes. In previous structural studies we found that an isolated difluorotoluene deoxynucleoside did not significantly distort a DNA duplex, even though when paired with adenine it significantly lowered the stability of the helix.[22] However, the F–Z pair might well be expected to be structurally different from an F–A pair, because analogue Z is larger than adenine by approximately 0.7 Å due to the presence of the H3 proton (Figure 1). In addition, any putative H-bonding between F and A would be lost. Thus some distortion of the DNA might be expected with the F–Z pair, base flipping being a distinct possibility.

To evaluate this question, we have carried out structural studies on a twelve base pair DNA duplex in which an F–Z base pair replaces a T–A base pair in the context of the consensus binding site of the human transcription factor SRY (Figure 1).[23] The structure was determined in aqueous solution by use of 2D-NMR methods combined with restrained molecular dynamics. We find that, despite the destabilization of the duplex by the F–Z pair relative to its natural counterpart, the predominant conformation is one in which the DNA is distorted very little. This finding is consistent with the idea that, despite their lack of Watson–Crick hydrogen bonds, such nonpolar base pairs are efficient and selective substrates for polymerase enzymes because they mimic the shape of natural pairs. The results also suggest that the F–Z pair might serve as a useful isosteric replacement for the T–A pair in probing H-bonded interactions with proteins and drugs, without fear of interference by a non-canonical structure.

Experimental Section

Sample Preparation

The 2,4-difluorotoluene[24] and 4-methylbenzimidazole[19] phosphoramidites were synthesized and incorporated into a DNA 12-mer using an Applied Biosystems 392 DNA synthesizer. The unmodified and modified oligonucleotides were synthesized in trityl-on mode and purified by HPLC.[25] The purity was checked by PAGE and found to be greater than 95%. The strands were mixed while monitoring the integrated intensity of the base protons in D_2O at 65 °C to obtain a 1:1 ratio. The sample was dissolved in buffer (100 mM NaCl, 1 mM EDTA, and 10 mM sodium phosphate, pH 7.4) at a duplex concentration of 3 mM.

Thermal Denaturation

The absorbance at 260 nm was monitored as a function of temperature (5–90 °C) on a Varian Cary One spectrophotometer. The duplexes were dissolved in 1 M NaCl, 10 mM sodium phosphate, and 0.1 mM EDTA at a pH of 7.0 to give duplex concentrations of 0.5, 1.25, 2.5, 3.75, 5, and 7.5 μM. All curves were fit assuming a two state model.[26,27] The thermodynamics of duplex formation were extracted by Van't Hoff analysis. The free energies derived from the individual curves agreed well with those from the Van't Hoff analysis, supporting a two-state model. Error in T_m and in free energy is estimated at ±0.5 °C and ±0.5 kcal/mol, respectively, on the basis of previous experience with this instrument and fitting method.

NMR Spectra

Experiments were performed on a Varian Unity Innova 500 MHz NMR spectrometer. Proton assignments were made using standard two-dimensional techniques, including NOESY, DQF-COSY, and NOESY WATERGATE. Data processing was done using Felix 97 (BIOSYM/Molecular Simulations). All structural restraints involving non-exchangeable protons were derived from NOESY data acquired at 15 °C with mixing times of 75, 150, 225, 300, and 375 ms. Restraints for exchangeable protons were derived from NOESY WATERGATE data acquired at 3 °C with a mixing time of 150 ms. Backbone conformation and sugar pucker were investigated at 15 °C using one-dimensional ^{31}P and DQF-COSY experiments, respectively. One-dimensional $^{1}H\{^{19}F\}$ heteronuclear NOE experiments were performed using a Varian triple resonance probe (^{15}N, ^{13}C, and ^{1}H). Experiments were performed in D_2O at 15 °C using the S2PUL pulse sequence, and in H_2O solution the BINOM pulse sequence was used with 1:1 water suppression at 3 °C. Because long irradiation times were employed in these experiments the NOEs observed were not incorporated into the restraint file.

Restraint Generation

The 206 restraints involving non-exchangeable protons were derived using MARDIGRAS.[28–30] From each experimental volume a distance was derived using an iterative relaxation matrix for each mixing time. One standard deviation was subtracted from the average lower bound and one standard deviation was added to the average upper bound calculated by MARDIGRAS to give the final upper and lower bounds. The Z19(H8) resonance was not observed in the spectra used for NOE build up as a result of exchange with D_2O solvent. Twenty-seven exchangeable proton restraints were given an upper bound of 5 Å and a lower bound of 1.5 Å. Dihedral angle restraints were used to preserve a right-handed DNA helix.[31–33] The δ torsion angle was restrained to $110° \leq \delta \leq 170°$ (C2′-endo sugar pucker) based on a greater than 14 Hz separation for the outer two peaks in the 1′ frequency axis of a DQF-COSY spectrum. A total of 233 NOE distance restraints and 132 dihedral angle restraints were used with 50 kcal mol^{-1} $Å^{-2}$ and 50 kcal mol^{-1} rad^{-2} energy terms, respectively, when applied full scale. In addition, 63 restraints were included to maintain Watson–Crick pairing for base pairs 1–4 and 8–12,[30] consistent with NOE data.

Molecular Dynamics

Starting structure generation and visualization of calculated structures was done using InsightII 95.0.6 (BIOSYM/Molecular Simulations). All structural calculations and analyses were done using the Discover, Analysis, and NMR Refine modules of InsightII. The AMBER force field in InsightII was modified to accommodate the fluorine atom using parameters from the AMBER95 force field in MACROMODEL v. 6.0 (C. Still, Columbia University). The partial charges for the 2,4 difluorotoluene nucleoside have been reported[22] while those for 4-methylbenzimidazole were assigned by ab initio methods at the RHF/6-31G(d) level using Gaussian 94 (Frisch, M. J. et al., Gaussian Inc.). The partial charges used in the molecular dynamics simulation for the 4-methylbenzimidazole monomer are the following; C1 = −0.25, H1 = 0.23 C2 = −0.21, H2 = 0.23, C3 = −0.27, H3 = 0.22, C4 = 0.17, C5 = 0.12, C6 = −0.01, C7Me = −0.64, C7MeH's = 0.23, N7 = −0.56, N9 = −0.07, C8 = 0.32, and H8 = 0.22. Molecular dynamics simulations consisted of an initial 5000 fs of equilibration followed by 10000 fs of dynamics starting with random velocities at 1000 K. The force constants were scaled from an initial scaling factor of 0.1 to a final factor of 1. The system was then equilibrated at 300 K for 5000 fs and subjected to 10000 fs of dynamics with a restraint scaling factor of 1. The third phase consisted of 5000 iterations of steepest descents energy minimization at 300 K with a restraint scaling factor of 1. Last, conjugate gradient energy minimization was run with a distance-dependent dielectric of 1 (the electrostatic terms were turned off in all previous phases) and a restraint scaling factor of 1 for 5000 iterations at 300 K. All phases of the

calculation were run with a Lennard-Jones potential with a cutoff of 20 Å. The global covalent scale and global nonbonded scale were set to 1 in each phase. Fifty calculations were performed using B-form DNA as the starting structure and a second set of 50 calculations were performed with an A-form starting structure. Ten of 50 final structures starting from B-form DNA and 3 of 50 final structures starting from A-form DNA contained severely collapsed minor grooves and were removed before averaging. The 40 and 47 remaining structures, respectively, were averaged and subjected to an additional 100 iterations of steepest descents energy minimization with full restraints to generate the final structures.

Results

Thermal Denaturation

The thermodynamics of duplex formation was investigated for the F–Z, F–A, and T–A duplexes by Van't Hoff analysis. The sequence context is the binding site for the transcription factor SRY in humans (Figure 1). The melting temperature was recorded by UV monitored thermal denaturation at six different concentrations (0.5–7.5 μM duplex) in a buffer containing 1 M NaCl, 10 mM phosphate and 0.1 mM EDTA pH 7.0. The T_m values at 5 μM duplex concentration for the F–A, F–Z, and T–A duplexes are 46.1, 49.8, and 58.2 °C, respectively (Figure 2). The free energies (37 °C) for the F–A, F–Z, and T–A duplexes are −10.0, −10.5, and −13.7 kcal/mol. Thus, the F–Z pair is destabilizing to the duplex relative to the natural pair, but appears to be somewhat less destabilizing than F–A.

The F–Z duplex is B-form

Structural studies were carried out at 500 MHz in aqueous buffer for the F–Z containing duplex. The chemical shifts of the imino protons for G•C and A•T base pairs (Figure 3) are consistent with formation of Watson–Crick base pairs. Cross-peaks in a NOESYWATERGATE spectrum recorded in H_2O solution confirm formation of G•C and A•T Watson Crick pairs. The T5 imino proton in the F–Z duplex is 0.43 ppm upfield of the corresponding T5 imino proton in the T–A duplex.[22] This upfield shift was also observed in a F–A duplex where the T5 imino proton was located 0.35 ppm upfield relative to the unmodified duplex.[22] In contrast, the G7 imino proton of the F–Z duplex shifts only 0.05 ppm upfield relative to the G7 imino proton of the unmodified duplex. A small downfield shift of 0.03 ppm was observed when comparing the G7 imino proton of the F–A duplex with that of the unmodified duplex.[22] With the exception of T5, all other imino proton chemical shifts in the F–Z duplex are within 0.1 ppm of the corresponding imino protons in the unmodified duplex. The temperature-dependent spectra of the imino proton region indicate broadening of the T5 imino proton resonance at lower temperature than other A–T base pairs in the F–Z duplex. Increased broadening suggests that the T5 imino proton has a shorter lifetime for base pair opening, although catalyzed exchange experiments are needed to quantify lifetimes. The transient opening of the F–Z pair is of interest but the lack of an exchangeable proton on either Z or F precludes determination of lifetimes by standard techniques. A similar temperature-dependent broadening of the T5 imino proton was seen in the duplex containing F–A.[22]

The non-exchangeable ^{1}H assignments for the F–Z duplex (Table 1) were made using standard techniques.[30] The base to H1′ (Figure 4), base to H2′/H2″, and base to H3′ regions show uninterrupted NOE connectivities throughout both strands of the F–Z duplex, indicative of a relatively undistorted DNA duplex. Uninterrupted NOE connectivities were also observed for the previously described T–A and F–A containing duplexes.[22] All observable base to H1′ intensities indicate that all bases adopt an *anti* conformation. Previously reported data showed that F adopts an *anti* conformation in solution, both in the nucleoside form[34] as well as within a DNA duplex.[22] The Z nucleoside was also found to adopt an *anti* conformation in solution.[19] Examination of the H1′ to H2′/H2″ region of the DQF-COSY spectra for the F–Z duplex at

15 °C indicates that all observable cross-peaks (including T5, G7, C18, and A20) exhibit a pattern and spacing similar to that of the control T–A duplex, confirming similar sugar puckers for all observable resonances. Furthermore, all resolvable H1′ to H2′/H2″ cross-peaks region show a spacing greater than 14 Hz between the outer two peaks of the H1′ frequency axis indicating an S(C2′-endo) sugar pucker.[31–33] Both the F6 and Z19 H1′ to H2′ H2″ cross-peaks are in overlap regions and therefore their sugar puckers were not restrained during molecular dynamics simulation. The non-exchangeable proton chemical shifts for the F–Z, F–A, and T–A duplexes are similar except for F and Z and their immediate neighbors. The similarity of the chemical shifts and DQF-COSY cross-peak patterns suggests the three duplexes adopt similar conformations. A ^{31}P NMR spectrum shows that all the phosphorus resonances reside within a one ppm range for the F–Z duplex, which is consistent with a B-form DNA structure.[35]

Both F6 and Z19 Are Stacked within the DNA Helix

Examination of a NOESY spectrum in D_2O at 15 °C shows that the A20(H2) proton exhibits NOEs to the adjacent F6(H3), Z19(H2), and Z19(H1) protons (Figure 4). NOEs are seen from the Z19(Me) protons to both the F6(H3) and C18(H6) protons. A further connectivity is seen from the F6(Me) protons to the T5(H6) proton. In a NOESYWATERGATE experiment at 3 °C both the T5 and G7 imino protons show NOEs to the F6(H3), Z19(H1), Z19(Me), and Z19 (H2) protons. These numerous contacts are consistent with a predominant conformation in which both F6 and Z19 are stacked within the DNA helix. The partial overlap of the Z19(H1) and F6(H3) resonances precluded the measurement of an NOE between them. The F6(H3) to Z19(H2) NOE was also not quantitated due to overlap of the Z19(H1′) and F6(H3) protons.

Heteronuclear ^{19}F NOEs further confirm the location of Z19 and F6 within the helix. The ^{19}F spectrum of the F–Z duplex shows two major peaks (data not shown). Upon irradiation of the F6(^{19}F4) resonance, which arises from the fluorine in position four of 2,4 difluorotoluene, a large NOE was seen to F6(Me) allowing the two fluorine nuclei to be assigned. The F6(^{19}F4) fluorine resonance also exhibited NOEs to both the Z19(Me) and the T5(Me) protons, confirming the location of F6(^{19}F4) in the major groove. The close proximity of the Z19(H1) chemical shift to the F6(H3) chemical shift prevented resolution of the F6(^{19}F4) to Z19(H1) NOE from that of the F6(^{19}F4) to F6(H3) NOE. Upon irradiation of the F6(^{19}F2) fluorine resonances, NOEs were seen to the Z19(H2) and the A20(H2) protons, confirming the location of F6(^{19}F2) in the minor groove.

Structural Calculations and Structural Features

NMR data sets were collected and analyzed to obtain 233 distance restraints and 132 dihedral angle restraints. Non-exchangeable proton restraints were obtained by inputting volumes from NOESY data at several different mixing times into MARDIGRAS.[28–30] The MARDIGRAS output contains both lower and upper bounds for each of the input volumes. The non-exchangeable proton restraints were combined with hydrogen bonding, dihedral angle, and exchangeable proton restraints to generate the input file used for molecular dynamics/energy minimization simulations. Both A-form and B-form 12-mer duplexes containing the F6-Z19 pair were used as initial structures in these simulations. The convergence of these structures is illustrated in Figure 5, and structural statistics are given in Table 2.

The final coordinates from forty molecular dynamics/energy minimization simulations were averaged and then subjected to 100 iterations of steepest descents minimization to give the final structure shown in Figure 6. The coordinates for the final structure have been deposited in the Protein Data Bank as entry 1EEK. In all calculated structures the F6 and Z19 are opposite one another within the DNA helix. Examination of the final structure shows that the F6-Z19 pair contains a slight propeller twist. The C1′ to C1′ distance increases 0.8 Å for the F6-Z19 pair, relative to the C1′ to C1′ distance found for normal B-form DNA, consistent with the

increased steric bulk associated with the Z19(H3) proton relative to adenine, which lacks this atom. The F6(H3) proton resides in the notch between the Z19(H3) and Z19(H2) protons which eliminates the overlap of the Z19(H3) and F6(H3) protons, depicted in Figure 1. This arrangement is 60° offset from the normal collinear arrangement of the N3•H3–N3 atoms found in a normal A–T base pair. The sugars of the F6 and Z19 bases both adopt S-type sugar puckers consistent with the conformational studies done for both the 2,4 difluorotoluene[34] and 4-methylbenzimidazole[19] nucleosides.

The F6(H3) proton within the F–Z duplex is found to resonate at 5.87 ppm at 20 °C. In the previously reported F–A structure[22] the F6(H3) proton was found to resonate at 6.38 ppm at 20 °C, significantly downfield of the F6(H3) proton of the F–Z duplex. By comparison, the H3 proton of 2,4 difluorotoluene was found to resonate at 6.49 ppm at 20 °C within the single strand, and 6.94 at 20 °C within the nucleoside. The temperature dependence of the F6(H3) proton was also of interest due to the 0.30 ppm downfield shift observed in the previously reported F–A duplex upon lowering the temperature from 30 to 3 °C. In contrast, the F6(H3) resonance of the F–Z duplex, the F containing single strand, and in the F nucleoside shows no significant temperature dependence. An explanation for the chemical shift differences and the temperature dependence of the H3 resonance of 2,4-difluorotoluene is not readily apparent and is most likely due to a combination of factors.

Discussion

The difluorotoluene–4-methylbenzimidazole pair is the first example of a nonpolar base pair surrogate for the T–A Watson–Crick base pair that contains a completely nonpolar pairing face. The F–Z pair has been shown to be replicated with good efficiency and fidelity by the Klenow fragment of DNA polymerase I,[3,4] which supports our recent hypothesis that steric matching, even in the absence of hydrogen bonds, may be a significant factor governing nucleotide insertion.[7,8,21] The structure shown in Figures 5 and 6 supports this hypothesis by showing minimal perturbation to the DNA helix, despite the addition of 0.7 Å to the Watson–Crick face. This suggests that an important factor in the ability of DNA polymerase to process the F–Z pair efficiently is due to its ability to discriminate matched from unmatched shapes, rejecting those that are not sterically accommodated within the enzyme active site.

The reduced thermal stability of the F–Z duplex when compared to the A–T duplex which is associated with pairing of Z with F might have led one to expect significant deviation from B-form structure, including the possibilities of a *syn* glycosidic orientation of either base, extrahelical location of either F or Z, extrahelical location of both F and Z, interdigitation of F and Z, and distortion of neighboring pairs. However, the F–Z duplex is clearly a B-form type helix with S-type sugar puckers and *anti* glycosidic angles for both the F6 and Z19 bases. In the predominant conformation the Z19 and F6 bases are both stacked within the DNA helix and the F6(H3) proton lies in the notch between the Z19(H2) and Z19(H3) protons. The only perturbation is a slight expansion of the C1′-to-C1′ distance which the enzyme appears to tolerate. The fact that the enzyme processes F–Z somewhat less efficiently than T–A may be due in part to this small distortion.

The resemblance of the Z–F pair to a natural A–T base pair (Figure 7) may make it a useful probe for hydrogen bonding and hydrophobic contacts within protein-DNA complexes. Most other nucleoside analogues used for this purpose have differed both in hydrogen bonding ability as well as steric size, thus making it difficult to decipher whether hydrogen bonding or local structural disruption is responsible for the observed changes. The use of the Z–F pair in conjunction with the F–A, T–Z, F–Q, and T–Q pairs (Q = Z with C1 replaced by N1 within the minor groove)[36] provides a new way to probe electrostatic and hydrophobic interactions in protein–DNA complexes.

Pyrimidines containing a fluorine nucleus at the C-5 position provide a probe nucleus located close to one side of the major groove, and have been used, for example, in studies of protein-nucleic acid complexes by NMR.[37] Use of the Z–F pair as an isosteric, nonpolar, A–T base pair surrogate provides two fluorine nuclei as NMR structural probes with one fluorine nucleus located in the center of both the major and minor grooves of DNA.

References

1. Schweitzer BA, Kool ET. J Org Chem 1994;59:7238–7242. [PubMed: 20882116]
2. Liu D, Moran S, Kool ET. Chem Biol 1997;4:919–926. [PubMed: 9427657]
3. Moran S, Ren RX, Kool ET. Proc Natl Acad Sci USA 1997;94:10506–10511. [PubMed: 9380669]
4. Moran S, Ren RXF, Rumney S IV, Kool ET. J Am Chem Soc 1997;119:2056–2057. [PubMed: 20737028]
5. Schweitzer BA, Kool ET. J Am Chem Soc 1995;117:1863–1872. [PubMed: 20882111]
6. (a) McMinn DL, Ogawa AK, Wu YQ, Liu JQ, Schultz PG, Romesberg FE. J Am Chem Soc 1999;121:11585–11586. (b) Ogawa AK, Wu YQ, McMinn DL, Liu JQ, Schultz PG, Romesberg FE. J Am Chem Soc 2000;122:3274–3287.
7. Kool ET. Biopolymers 1998;48:3–17. [PubMed: 9846123]
8. Kool ET, Morales JC, Guckian KM. Angew Chem, Int Ed 2000;39:990–1009.
9. Evans TA, Seddon KR. Chem Commun 1997:2023–2024.
10. Howard JAK, Hoy VJ, O'Hagan D, Smith GT. Tetrahedron 1996;52:12613–12622.
11. Thalladi VR, Weiss HC, Bláser D, Boese R, Nangia A, Desiraju GR. J Am Chem Soc 1998;120:8702–8710.
12. Plenio H, Diodone R. Chem Ber 1997;130:633–640.
13. Meyer M, Suhnel J. J Biomol Struct Dyn 1997;15:619–624. [PubMed: 9440008]
14. Wang X, Houk KN. Chem Commun 1998:2631–2632.
15. Diederichsen U. Angew Chem, Int Ed 1998;37:1655–1657.
16. Barsky D, Colvin M, Kool ET. J Biomol Struct Dyn 1999;16:1119–1134. [PubMed: 10447197]
17. Cubero E, Sherer EC, Luque FJ, Orozco M, Laughton CA. J Am Chem Soc 1999;121:8653–8654.
18. Schmidt KS, Sigel RKO, Filippov DV, van der Marel GA, Lippert B, Reedijk J. New J Chem 2000;24:195–197.
19. Guckian KM, Morales JC, Kool ET. J Org Chem 1998;63:9652–9656. [PubMed: 20852720]
20. Seela F, Bourgeois W, Rosemeyer H, Wenzel T. Helv Chim Acta 1996;79:488–498.
21. Morales JC, Kool ET. Nat Struct Biol 1998;5:950–954. [PubMed: 9808038]
22. Guckian KM, Krugh TR, Kool ET. Nat Struct Biol 1998;5:954–959. [PubMed: 9808039]
23. Harley VR, Lovell-Badge R, Goodfellow PN. Nucleic Acids Res 1994;22:1500–1501. [PubMed: 8190643]
24. Chaudhuri NC, Kool ET. Tetrahedron Lett 1995;36:1795–1798.
25. Huang G, Krugh TR. Anal Biochem 1990;190:21–25. [PubMed: 2285142]
26. Petersheim M, Turner DH. Biochemistry 1983;22:256–263. [PubMed: 6824629]
27. Freier SM, Burger BJ, Alkema D, Neilson T, Turner DH. Biochemistry 1983;22:6198–6206.
28. James TL. Curr Opin Struct Biol 1991;1:1042–1053.
29. Borgias BA, James TL. J Magn Reson 1990;87:475–487.
30. Schmitz U, James TL. Methods Enzymol 1995;261:3–44. [PubMed: 8569500]
31. Baleja JD, Pon RT, Sykes BD. Biochemistry 1990;29:4828–4839. [PubMed: 2141998]
32. Gronenborn AM, Clore GM. Biochemistry 1989;28:5978–5984. [PubMed: 2550066]
33. Clore GM, Oschkinat H, McLaughlin LW, Benseler F, Happ CS, Happ E, Gronenborn AM. Biochemistry 1988;27:4185–4197. [PubMed: 3415980]
34. Guckian KM, Kool ET. Angew Chem, Intl Ed Engl 1997;36:2825–2828.
35. Gorenstein DG. Methods Enzymol 1992;211:254–286. [PubMed: 1406310]

36. Morales JC, Kool ET. J Am Chem Soc 1999;121:2323–2324. [PubMed: 20852718]
37. Rastinejad F, Evilia C, Lu P. Methods Enzymol 1995;261:560–576. [PubMed: 8569512]

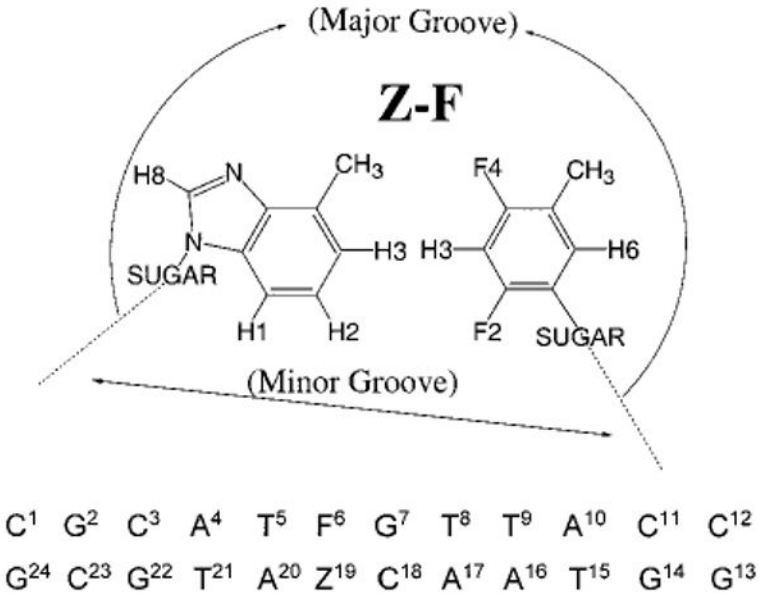

Figure 1.
(A) Schematic model of a 4-methylbenzimidazole opposite a 2,4 difluorotoluene (Z–F) pair. (B) Sequence of the 12-mer duplex containing the Z–F pair in the center of the consensus binding sequence for the SRY protein.[21]

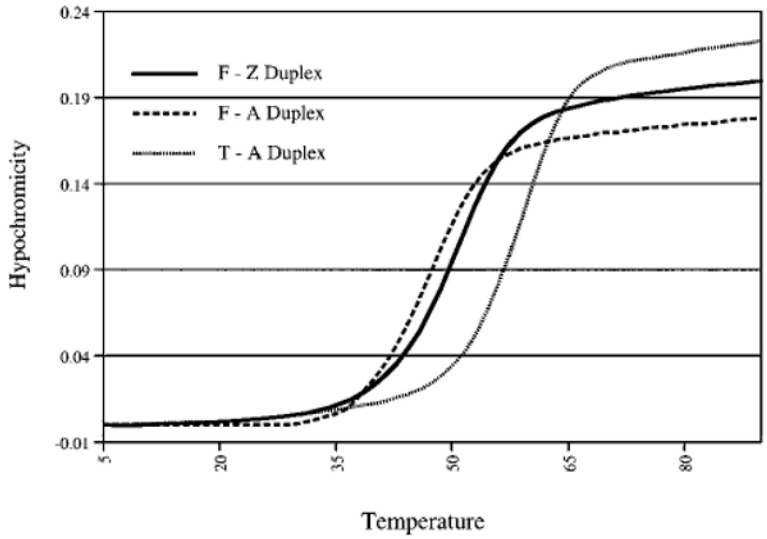

Figure 2.

Thermal denaturation curves at 5 μM duplex concentration in a buffer containing 1 M NaCl, 10 mM sodium phosphate (pH 7.0), and 0.1 mM EDTA for the F–Z, F–A, and T–A duplexes. The melting temperatures at 5 μM duplex concentration for the duplexes are 46.1, 49.8, and 58.2 °C, respectively.

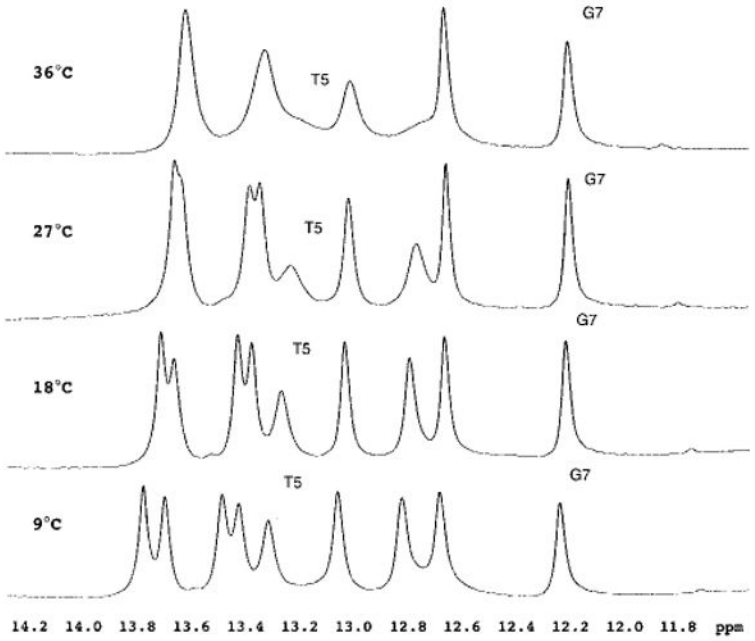

Figure 3.
Imino proton spectra at 500 MHz for the F–Z duplex at 9, 18, 27, and 36 °C in H_2O. The terminal imino protons are visible at 3 °C, but disappear by 9 °C due to exchange with water.

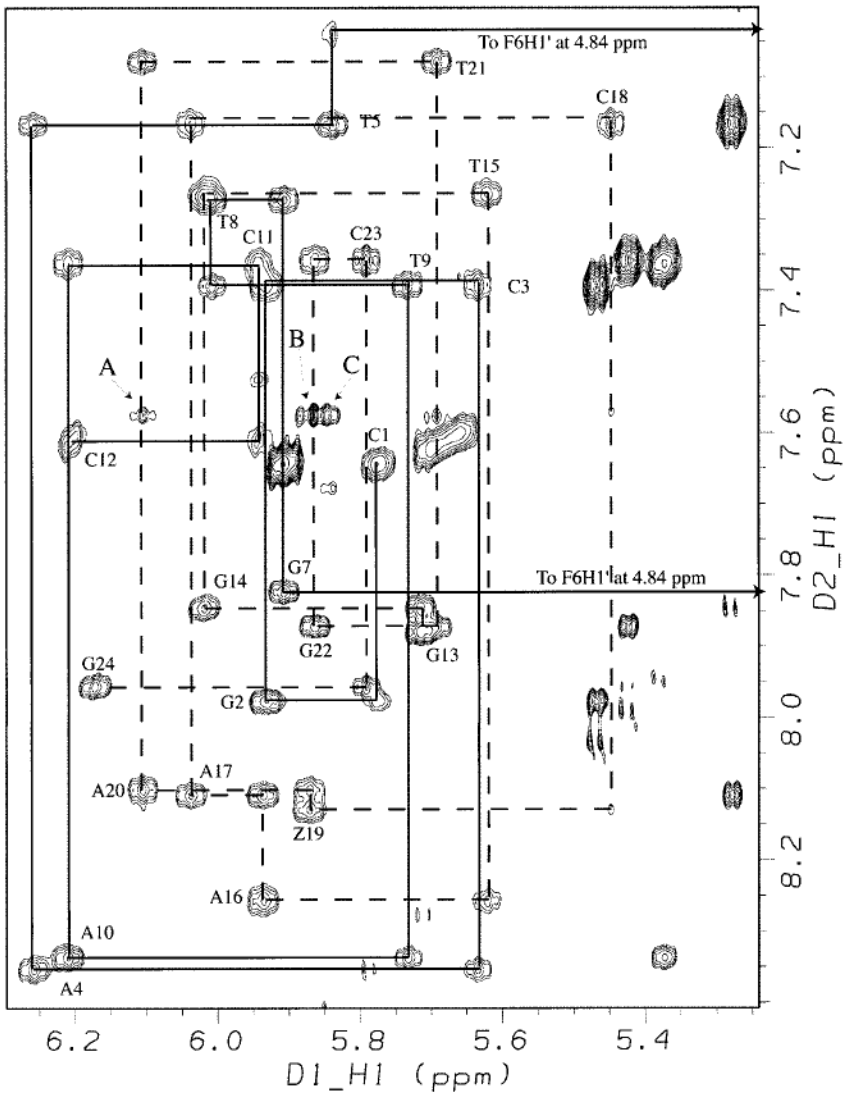

Figure 4.

Base to H1′ region of a 375 ms NOESY spectrum taken at 15 °C. Labeled cross-peaks are: A, A20(H2) to Z19(H2); B, A20(H2) to F6(H3); C, A20(H2) to Z19(H3).

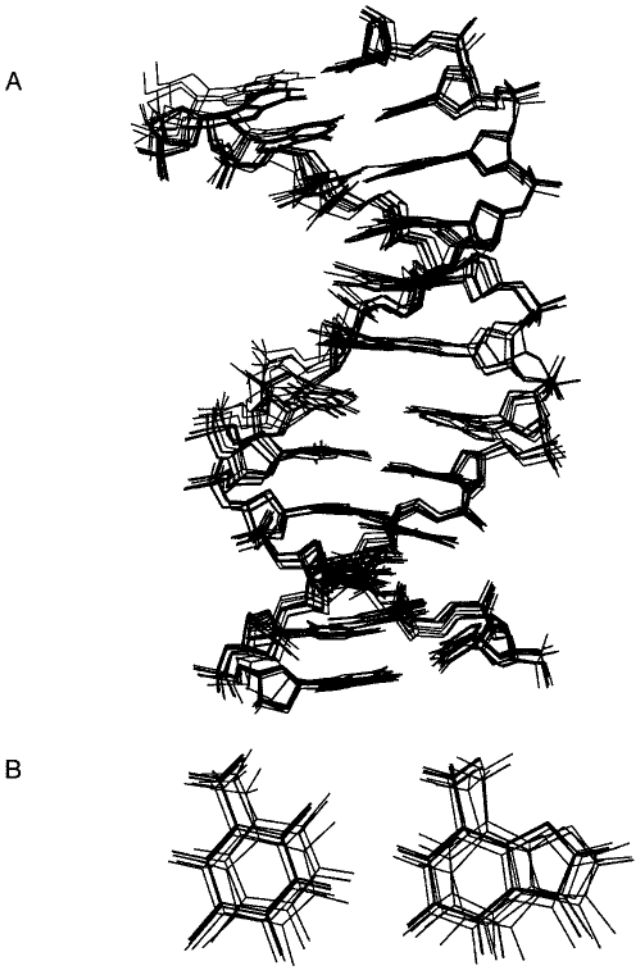

Figure 5.
(A) Superimposition of 12 randomly selected final structures after restrained molecular dynamics and energy minimization. Six were started from B-form DNA while six were started from A-form DNA. (B) Overlay of the F–Z pairs from the 12 structures shown in (A).

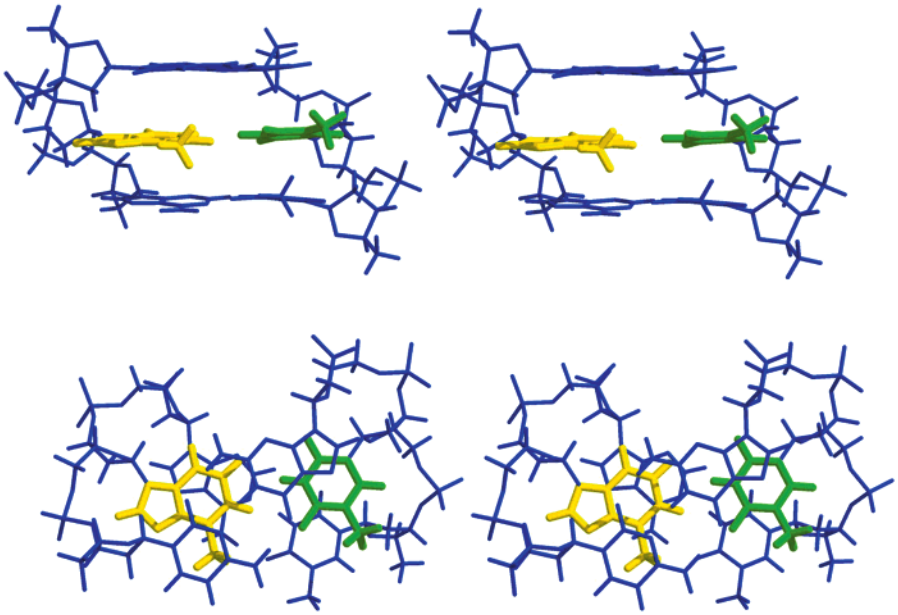

Figure 6.
Stereoviews of the central three base pairs for the Z–F duplex. (A) View from the major groove, showing the F6 (yellow) and Z19 (green) pair as well as their immediate neighbors (blue). (B) Cutaway view along the helical axis. Figures were produced using InsightII (Molecular Simulations Inc.).

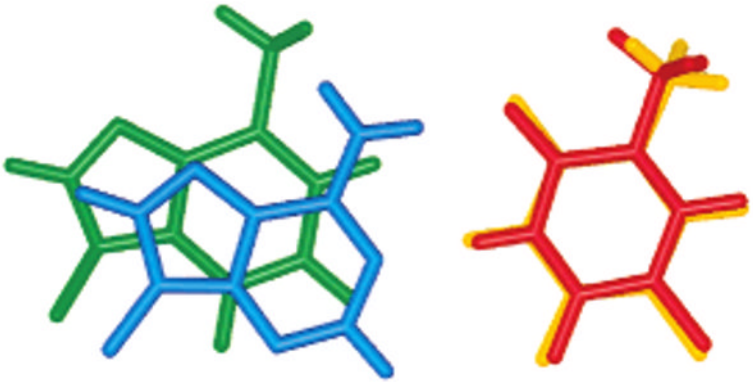

Figure 7.
Overlay of the Z–F pair from the final structure with an A–T pair. The Z–F pair is shown in yellow (F) and green (Z) while the A–T pair is shown in red (T) and blue (A).

Table 1

Proton Chemical Shifts (ppm) for the Z–F Duplex[a]

base	H8/H6	H5/Me/H2	H1′	H2′	H2″	H3′	H4′	H1/H3	NH_b/NH_{nb}[b]
C1	7.65	5.90	5.77	1.99	2.43	4.72	4.38		*c*
G2	7.98		5.93	2.70	2.77	5.00	4.38	13.07	
C3	7.40	5.46	5.64	2.06	2.42	4.88	4.38		6.56/8.38
A4	8.36	7.64	6.26	2.70	2.90	5.05	4.45		
T5	7.17	1.53	5.84[d]	1.72	2.22	4.81[d]	4.42	13.35	
F6	7.04	1.76	4.84	2.01	2.22	4.78	4.38	5.86	
G7	7.83		5.91	2.65	2.76	4.95	4.38	12.25	
T8	7.27	1.29	6.02	2.13	2.54	4.85	4.38	13.82	
T9	7.40	1.66	5.73	2.12	2.47	4.90[e]	4.46	13.71	
A10	8.34	7.47	6.21	2.75	2.87	5.05	4.45		
C11	7.37	5.37	5.94	2.42	2.08	4.70	4.45		6.75/8.22
C12	7.62	5.66	6.21	2.26	2.26	4.54	4.31		*c*
G13	7.89		5.71	2.58	2.71	4.83	4.41	c	
G14	7.85		6.02	2.65	2.80	4.99	4.41	12.84	
T15	7.27	1.46	5.62	2.07	2.39	4.90[e]	4.42	13.52	
A16	8.26	6.99	5.94	2.75	2.89	5.08	4.42		
A17	8.11	7.53	6.03	2.54	2.78	5.03	4.43		
C18	7.17	5.27	5.45	1.51	1.89	4.68	4.42		6.89/8.00
Z19	8.14	1.83/6.09	5.87	2.61	2.72	4.98	4.44	6.59/5.84	
A20	8.11	7.58	6.11	2.62	2.86	4.99	4.44		
T21	7.08	1.30	5.69	1.99	2.37	4.85	4.38	13.45	
G22	7.88		5.86	2.62	2.70	4.98	4.38	12.70	
C23	7.36	5.42	5.79	1.93	2.36	4.83	4.38		6.67/8.44
G24	7.96		6.17	2.64	2.39	4.69	4.43	c	

[a] Non-exchangeable proton chemical shifts were measured at 15 °C and referenced to the HDO signal at 4.90 ppm, the exchangeable proton chemical shifts were measured at 3 °C and referenced to the HDO signal at 5.05 ppm.

[b] The abbreviations NH_b and NH_{nb} refer to the hydrogen-bonded and non-hydrogen-bonded amino protons in cytosine.

[c] Chemical shift not determined due to exchange broadening.

[d] Assignment made 40 °C.

[e] Assignment made at 20 °C.

Table 2

Structural Statistics for the F–Z Containing Duplex

Structural Restraints		
distance restraints	total	296
	intraresidue	94
	interresidue	202
	exchangeable	27
	non-exchangeable	206
	hydrogen-bonding	132
dihedral angle restraints		132
Violations of Experimental Restraints in the Final Structure When:		
A-form DNA was used as a starting structure:	distance violations(.0.1 Å)	<SA>[a]
	total	2[c]
	intraresidue	0
	interresidue	2
	dihedral violations (>2°)	0
B-form DNA was used a a starting structure:	distance violations (.0.1 Å)	<SA>[b]
	total	1[c]
	intraresidue	0
	interresidue	1
	ihedral violations (>2°)	1
Atomic rms Differences (Å):[d]		
final (A-form starting) and final (B-form starting		0.79
Final (B-form starting) and B-form starting structure		2.69

[a] Energy-minimized average structure from 47 calculations.

[b] Energy-minimized average structure from 40 calculations.

[c] No violations >0.2 Å

[d] Average pairwise rms differences were calculated using energy-minimized average structures.