Solution Structure of a Five-Adenine Bulge Loop within a DNA Duplex†

Utz Dornberger, Alexander Hillisch,‡ Friedrich A. Gollmick, Hartmut Fritzsche, and Stephan Diekmann*,‡

Institut für Molekularbiologie, Friedrich-Schiller-Universität, Winzerlaer Strasse 10, D-07745 Jena and Institut für molekulare Biotechnologie, Beutenbergstrasse 11, D-07745 Jena, Germany

Received March 24, 1999; Revised Manuscript Received June 9, 1999

ABSTRACT: The three-dimensional solution structure of a DNA molecule of the sequence 5′-d(GCATC-G**AAAAA**GCTACG)-3′ paired with 5′-d(CGTAGCCGATGC)-3′ containing a five-adenine bulge loop (dA_5-bulge) between two double helical stems was determined by 2D ^{1}H and ^{31}P NMR, infrared, and Raman spectroscopy. The DNA in both stems adopt a classical B-form double helical structure with Watson−Crick base pairing and C2′-endo sugar conformation. In addition, the two dG/dC base pairs framing the dA_5-bulge loop are formed and are stable at least up to 30 °C. The five adenine bases of the bulge loop are localized at intrahelical positions within the double helical stems. Stacking on the double helical stem is continued for the first four 5′-adenines in the bulge loop. The total rise (the height) of these four stacked adenines roughly equals the diameter of the double helical stem. The stacking interactions are broken between the last of these four 5′-adenines and the fifth loop adenine at the 3′-end. This 3′-adenine partially stacks on the other stem. The angle between the base planes of the two nonstacking adenines (A10 and A11) in the bulge loop reflects the kinking angle of the global DNA structure. The neighboring cytosines opposite the dA_5-bulge (being parts of the bulge flanking base pairs) do not stack on one another. This disruption of stacking is characterized by a partial shearing of these bases, such that certain sequential NOEs for this base step are preserved. In the base step opposite the loop, an extraordinary hydrogen bond is observed between the phosphate backbone of the 5′-dC and the amino proton of the 3′-dC in about two-thirds of the conformers. This hydrogen bond probably contributes to stabilizing the global DNA structure. The dA_5-bulge induces a local kink into the DNA molecule of about 73° (±11°). This kinking angle and the mutual orientation of the two double helical stems agree well with results from fluorescence resonance energy transfer measurements of single- and double-bulge DNA molecules.

Base bulges are formed by nucleic acids when a duplex section is interrupted by one or more formally single-stranded bases on one strand that are unopposed by bases on the other strand. Base bulges in nucleic acids are an important feature in the repertoire of folding elements. They can be exploited for specific recognition of proteins.

DNA bulges may be created in heteroduplex DNA arising from recombination between imperfectly homologous sequences or from errors of replication. They become features to be recognized by the repair machinery. Bulged bases resulting from replicative errors are considered to play an important role in frame-shift mutagenesis (*1*). Certain intercalative mutagens have been found to bind with an increased affinity at or near bulge sites in DNA (*2*−*4*). RNA bulges occur frequently in the secondary structures of large RNAs. Many bulge loops are phylogenetically conserved, indicating functional importance. In some cases, they participate in tertiary interactions (*5, 6*) or specific interactions with RNA binding proteins (*7*−*12*).

A number of different biophysical methods have been applied for the study of bulged nucleic acids. The data from gel electrophoresis (*13*−*17*), fluorescence resonance energy transfer (FRET)[1] measurements (*18, 19*), and cryo-electron microscopy (*20*) show a general agreement on the global shape of DNA and RNA molecules containing base bulges: the bulges introduce a defined kink into the helical axis. Several NMR and crystallographic studies of single-nucleotide bulges in RNA (*21*) and DNA (*22*−*26*) were performed to investigate the bulge structures at a more detailed stereochemical level. These studies indicate that in most cases the unpaired base is stacked within the helix, although pyrimidine base bulges within dAT-rich regions were found to be extrahelical (*22, 23*). The unpaired bases in three-nucleotide DNA bulges were found to be stacked into the helix continuously with the flanking DNA (*27, 28*), whereas in three-nucleotide RNA bulges of 5′UCU3′ (*11, 12*) from the trans-activation response (TAR) element of HIV, the 3′ uracil is partially or transiently extrahelical. It was suggested that a further extension of the unpaired region to five or six

† This study was supported by fonds of the Chemical Industry (H.F.), the Deutsche Forschungsgemeinschaft, the Deutsche Akademische Austauschdienst, and the Thuringian Ministry of Science, Research, and Culture.

* To whom correspondence should be addressed. Phone: +49-3641-656262. Fax: +49-3641-656272. E-mail: diekmann@imb-jena.de.

‡ Institut für molekulare Biotechnologie.

[1] Abbreviations: 2D, two-dimensional; 2QF-COSY, double quantum filtered correlation spectroscopy; FRET, fluorescence resonance energy transfer; FT-IR, Fourier transform infrared; NMR, nuclear magnetic resonance; NOE, nuclear Overhauser enhancement; NOESY, NOE spectroscopy; ppm, parts per million; rMD, restrained molecular dynamics; RMSD, root-mean-square deviation; TAD, torsion angle dynamics.

10.1021/bi9906874 CCC: $18.00
Published on Web 09/08/1999

nucleotides should enhance the flexibility and conformational complexity of DNA or RNA bulges (*13*). In the recently reported crystal structure of a five-nucleotide adenosine-rich RNA bulge in the group I intron from *Tetrahymena thermophila* (*6*), all the bulge nucleotides reside outside of the helix. In contrast to these observations, a corresponding NMR study (*29*) shows continuous stacking between the two 5′ and 3′ adenines in the bulge and the base-paired nucleotides flanking the bulge. Less information is available about the conformations of larger DNA bulges. Chemical and enzymatic probing of a five-nucleotide DNA bulge (*13*) suggested conformational complexity and diversity for large DNA bulges. Furthermore, FRET measurements (*18*) indicate that five-adenine bulge loops induce kinks of 85−105°. No high-resolution structures are yet available of DNA bulges with more than three unpaired nucleotides.

In our work, we used 2D ^{1}H and ^{31}P NMR spectroscopy and torsion angle dynamics calculations to describe the solution structure of a five-nucleotide DNA bulge containing five adenines (dA_5-bulge) in the unpaired region. We present a detailed analysis of the helical parameters and the kinking angle of the dA_5-bulge. Furthermore, we have used NMR spectroscopy to determine the base-pair dynamics of the base pairs flanking the five-adenosine bulge.

The overall DNA kinking angle induced by the dA_5-bulge obtained by these NMR data agrees with that obtained by a fluorescence resonance energy transfer (FRET) analysis (*30*).

EXPERIMENTAL PROCEDURES

Sample Preparation. The oligonucleotides were synthesized by using automated phosphoramidite chemistry on a DNA synthesizer (Expedite). Phosphoramidites and the standard solutions for activation, oxidation, and detritylation were purchased from GLEN Research. Supports consisted of a thin layer (3−5%) of polystyrene grafted onto a poly(tetrafluoroethylene) core (*31*). Polymer-supported oligonucleotides were cleaved from supports and deprotected by treatment with 28% aqueous ammonia solution for 6−12 h at 55 °C. Purification was carried out on a HPLC LC-10AT Shimadzu liquid chromatograph using a mono Q HR5/5 anion-exchange column from Pharmacia. The oligonucleotides were desalted on a Sephadex G10 column.

The NMR sample was prepared by combining equimolar amounts of the heptadecadeoxynucleotide 5′-d(GCATCGAAAAAGCTACG)-3′ and the dodecadeoxynucleotide 5′-d(CGTAGCCGATGC)-3′ in a phosphate buffer containing 100 mM NaCl and 0.05 mM EDTA at pH 7.0. The sample was lyophilized three times from D_2O and finally dissolved in 550 μL of 99.996% D_2O. The sample concentration was 2.8 mM in duplex. For assignments of exchangeable protons, a DNA sample in 10% D_2O and 90% H_2O was prepared. The pH value was checked in the NMR tube.

Vibrational Spectroscopy. FT-IR measurements were performed using a Bruker IFS-66 FT-IR spectrometer with DGTS detector. About 1.5 μL of the DNA solution at 7 mM duplex concentration in D_2O were deposited in cells equipped with ZnSe windows. Aquisition parameters were used as previously described (*32*).

The Raman spectra were acquired using a Bruker IFS-66 FT-IR spectrometer equipped with the Raman module FRA 106. About 5 μL of the DNA solution at 6 mM duplex concentration were filled in glass capillaries. The sample was excited with 514.5 nm lines of an argon ion laser (300 mW of radiant power at the sample). All experiments were carried out at room temperature. The Raman spectra were smoothed applying a Savitzky-Golay algorithm with 11 points.

NMR Experiments. ^{1}H and ^{31}P NMR experiments were run on a 500 MHz Varian Unity INOVA and a 600 MHz Varian UnityPlus spectrometer. All 2D measurements were carried out at 27 °C unless indicated otherwise.

Imino proton spectra were obtained by using the solvent suppression 1331 pulse sequence (*33*) with the carrier set at the H_2O resonance. Data were collected for a spectral width of 12 kHz. The temperature was varied between 5 and 50 °C.

2D NOE spectra with mixing times of 80, 140, 200, and 400 ms were recorded in D_2O. A spectral width of 5000 Hz was used and the carrier frequency set to the HDO resonance frequency. A total of 630 FIDs of 4K complex data points were collected in t_1 with 32 scans and a repetition delay of 6 s. NOESY data sets were multiplied with a Gaussian window function in both dimensions and zero-filled to give a final 4K × 4K matrix.

A pure absorption double-quantum-filtered COSY spectrum (*34*) was recorded using 4K complex points in the aquisition dimension, 600 experiments in the t_1 dimension with 90 scans at each t_1 value and a spectral width of 4500 Hz. After apodization in both dimensions with a squared sine bell shifted by 25°, zero filling was used to obtain a final data set of 8K × 1K corresponding to a digital resolution of 0.55 Hz/point in F2 and 4.4 Hz/point in F1.

A proton-detected ^{31}P-^{1}H heteronuclear correlation spectrum was collected with a spectral width of 4500 Hz in the ^{1}H dimension and 1000 Hz in the ^{31}P dimension: 2048 complex points in the t_2 (^{1}H) dimension and 192 complex points in the t_1 (^{31}P) dimension with 128 scans at each t_1 value were collected (*35*).

Measurements of the spin−lattice relaxation time T_1 was performed by the inversion−recovery method. Spin−spin relaxation times T_2 were determined using the Hahn spin−echo method. The isotropic correlation time $\tau_c = 1.8$ ns was calculated by the equation $\tau_c = 2\omega^{-1}(3T_2/T_1)^{-1/2}$ assuming a simple and rigid molecular motion (*36*).

One-dimensional spectra were obtained applying the inversion recovery pulse sequence to acquire partially recovered spectra in order to selectively observe adenine H2 resonances (*37*). A total of 8192 complex points were collected in a spectral width of 5000 Hz. The line widths were measured using the *dres* command in the VNMR (Varian Associates Inc.) software package or by fitting the spectra applying the FELIX 95 software (Biosym Inc.).

Generation of Interproton Distance Restraints. 2D NOE intensities were determined by the fitting method using a Gaussian function in the software package SPARKY, version 3.28 (*38*). Those cross-peaks which differed significantly in their intensity on both sides of the diagonal (>50%) were discarded from the NOE intensity set.

Interproton distances were calculated from the extracted NOE intensities and three isotropic correlation times (τ_c = 1.8, 2.5, and 3.0 ns) using the relaxation matrix software MARDIGRAS, version 3.2 (*39*). A three-site jump model was used to calculate rapidly rotating methyl protons. The initial model structure for the MARDIGRAS calculations was

generated applying a systematic conformational search in helicoidal space (*40*) using the program JUMNA, version 10.0 (*41*). A total of 486 starting geometries were specified by varying the helicoidal parameters that define the orientation of the unpaired bases and the two stems with respect to one another. These structures were subjected to an energy minimization applying the AMBER force field (*42*) in JUMNA. The 20 conformers (RMSD = 2.3 Å) with the lowest energy were averaged. A total of nine interproton distance data sets (three mixing times × three isotropic correlation times) were obtained for the dA_5-bulge molecule. For any particular proton−proton distance, the average value d and standard deviation Δd from the nine distance sets were calculated. The upper and lower limits of the distance restraints were set to $d \pm \Delta d$, but a minimum well width of 0.6, 0.4, or 0.2 Å was maintained in case of 1−3, 3−6, or 6−9 available distances, respectively. Strong NOEs involving exchangeable hydrogens observed in the 200 ms NOESY experiment were given an upper limit of 4.5 Å.

Additionally, 26 distance restraints defining the Watson−Crick hydrogen bonds of the nonterminal residues of the DNA duplex were included in the restraint set (*43*). Lower limit restraint of 9.0 Å between N1 and N9 of the bases involved in hydrogen bonding were introduced in the DYANA calculations to prevent base intercalation.

Determination of Torsion Angle Restraints. The determination of the sugar, backbone, and glycosidic torsion angles were carried out as previously described (*6, 44, 45*). The sugar conformations were characterized by the H1′-H2′ scalar couplings evident in the COSY experiments. The deoxyribose pseudorotational phase angle were constrained to 90−190° and to 110−180° for all sugars showing $^3J_{H1'-H2'}$ coupling constants of >8.8 and >10.5 Hz, respectively. The puckering amplitude was set to 36° for all calculations of the sugar torsion angles. The sugar conformations were further inspected by the observation of H2′2-H3′ and H3′-H4′ couplings.

The glycosidic torsion angle χ was constrained based on the examination of the H8/H6-H1′/H2′/H3′ distances derived from the 2D NOE spectra and MARDIGRAS calculations (*44, 46, 47*). The β and γ torsion angles were constrained following a qualitative interpretation of *J*-coupling data using the long-range (*n*)P−(*n*)H4′ four-bonds coupling (*45, 48*). Further restriction of the torsion angles β and γ was obtained by the inspection of line widths of the H4′, H5′1, and H5′2 proton resonances (*44*). The ϵ torsion angle was constrained by the measurement of the heteronuclear $^3J_{H3'-P}$ coupling constants and the observation that no long range $^4J_{H2'-P}$ was found in the ^{1}H-^{31}P correlation spectrum (*45, 49*). Additionally, the α and ζ torsion angles were restrained based on the observation of the ^{31}P chemical shifts (*50*).

No artificial angle restraints between the three atoms defining a hydrogen bond were included in the restraints set to permit the highest flexibility in orientation between complementary bases in a Watson−Crick base pair.

Structure Calculation. The initial distance geometry calculations were performed with the program DIANA (*51*). After the first round of structure calculations with the full restraints data set, several distance limit violations of >0.2 Å were observed. All distance limit violations were located in the double helical stems of the DNA molecule mainly involving the base/sugar to methyl NOEs and the sequential H6/H8 to H2″ NOEs. In a following step, these distance limits were adjusted by subtraction of maximal 0.5 Å (lower limits) or addition of maximal 0.5 Å (upper limits) to avoid distance limit violations > 0.2 Å. After the initial structure calculations, we also observed several close hydrogen−hydrogen contacts in the bulge loop region which were not defined by distance restraints. If no NOE cross-peaks were found for these hydrogen−hydrogen distances in the 200 and 400 ms NOE spectra, these distances were introduced as "non-NOE" distance restraints with a lower limit of 4.5 Å.

This data set was subsequently used in the torsion angle dynamics (TAD) calculations applying the software package DYANA 1.5 (*52*). Structure calculations were performed using the standard simulated annealing protocol with N = 20 000 TAD steps.

The 15 conformers with the smallest target function values were subjected to restrained energy minimizations in AMBER 4.1. (*53*). The energy optimizations were carried out in vacuo using the AMBER91 force field (*42*). The distance dependent dielectric function of $\epsilon = 4|r|$ (where r is the distance between interacting atoms, value of r in angstroms) was used to approximately simulate the surrounding aqueous solution. A residue-based cutoff of 12 Å was applied for the treatment of nonbonded interactions. In addition, whenever 25 optimization steps were performed all nonbonded interactions with an 18 Å cutoff were included. The same set of NMR-derived restraints which was applied in the distance geometry calculations was used in the restraint energy minimization. The AMBER potentials for restraints had the form of a deep well, consisting of a flat bottom with parabolic sides out to 0.1 Å and 5° for distance and torsion angle restraints, respectively, and linear sides beyond that. The force constants for distance and angle restraints were set to 20 kcal mol^{-1} Å^{-1} and 60 kcal mol^{-1} rad^2, respectively. During the first 100 optimization steps, the steepest descent minimizer was employed. The minimization was switched to the conjugate gradient method until the gradient norm of successive steps was below 0.05 kcal mol^{-1} Å^{-1}.

Structural Analysis. Analysis of the conformational and helical parameters was carried out with the programs MOLMOL, version 2.6 (*54*), and CURVES, version 4.1 (*55*). The local helical parameters, defined relative to a linear global helix axis, were extracted.

The kinking angle between the two stems was calculated with the program ARC-FIT (P. Slickers, personal communication). In a first step, the helical axis of each stem was calculated using CURVES. Then ARC-FIT was employed to fit a kinked line to both helical segments simultaneously. This procedure ensures that all bases in the stems are considered when calculating the kinking angle between the two stems. Thus, the kink angle is the angle between two lines giving the best match to the helical axes of the two stems.

RESULTS AND DISCUSSION

A DNA molecule containing five unpaired dAs (dA_5-bulge) between two double-helical stems of the sequence 5′-d(GCATCG)-3′ and 5′-d(GCTACG)-3′ (see Scheme 1) was analyzed by NMR, Fourier transform infrared (FT-IR), and Raman (FT-Raman) spectroscopy. Residues of the dA_5-bulge are numbered as given in Scheme 1.

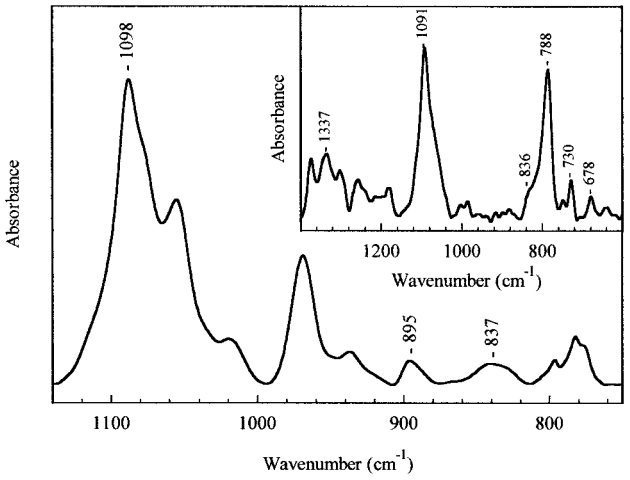
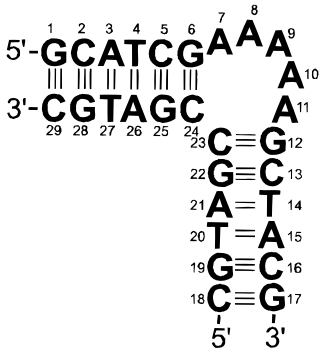


FIGURE 1: FT-infrared and FT-Raman (insert) spectrum of the dA_5-bulge collected at room temperature.

Scheme 1



Vibrational Spectroscopy. FT-IR and FT-Raman spectroscopy are very useful spectroscopic techniques for rapid evaluation of certain structural features of nucleic acids. The sugar-backbone conformation can be examined using vibrational marker bands. The FT-IR spectrum and FT-Raman spectrum of the dA_5-bulge in the sugar−phosphate region below 1300 cm^{-1} show the typical conformational marker bands observed for a B-DNA-like backbone conformation (see Figure 1). The infrared band around 837 cm^{-1} is indicative of C2′-endo sugar conformation. Additionally, the absence of a band around 860 cm^{-1} indicates that all sugars in the dA_5-bulge have predominantly C2′-endo conformation (*56*). This conclusion is further supported by a Raman band at 836 cm^{-1} which is an indicator for C2′-endo conformation (*57*). The presence of the adenine breathing mode band at 730 cm^{-1} in the Raman spectrum together with the absence of a vibrational band around 620 cm^{-1} (Figure 1, insert) confirms the anti conformation of the adenine bases in the dA_5-bulge (*57*).

^{1}H NMR Measurements. (1) Imino Proton Studies. The imino protons of the guanine and thymine bases in the dA_5-bulge molecule are observed in the low-field region between 12 and 14 ppm (see Figure 2). The low-field chemical shifts are characteristics of hydrogen-bonded imino protons in Watson−Crick base pairs (*58*) and are direct evidence for the formation of the double-helical structure. Thymine and guanine imino protons were assigned on the basis of the observation of strong NOEs to the nonexchangeable H2 proton of the adenine residue and the amino protons (hydrogen-bonded and exposed) of the cytosine residue,

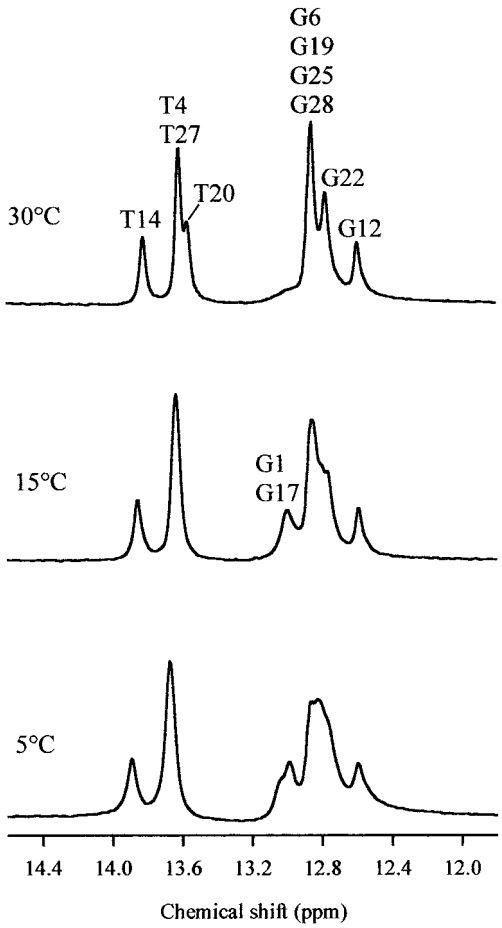


FIGURE 2: ^{1}H NMR imino proton spectra of the dA_5-bulge with assignment according to numbering in Scheme 1.

respectively, with which it forms a Watson−Crick hydrogen-bonded pair. The spectral presence of the G6 and G12 (see Figure 2) imino protons establishes that, in each duplex part, the G6/C24 and G12/C23 Watson−Crick base pairs are formed under the conditions studied, that is even in the presence of a five-adenine bulge loop between these base pairs. Furthermore, the guanine imino protons of G6 and G12 show NOEs to the amino protons (hydrogen-bonded and exposed) of the cytosine residue C24 and C23. All imino protons from nonterminal residues were observed even at 30 °C, indicating stable Watson-Crick base-pair formations in both duplex parts under the studied conditions. This result is consistent with earlier NMR data for bulged DNA molecules containing three unopposed bases (*59*). Aboul-ela et al. (*28*) also investigated a three-adenine DNA bulge. They however found that the presence of this bulge loop disrupts neighboring dG/dC as well as dA/dT base pairs on either side. They discussed that the extent of disruption of neighboring base pairs depends on the sequence context of the bulge loop. Our results show that the insertion of a five-adenine bulge loop in the DNA duplex does not result in the disruption of the flanking dG/dC base pairs.

The base pair dynamics of the base pairs flanking the dA_5-bulge loop were derived from ^{1}H NMR measurements of the imino proton exchange rates upon titration with exchange catalyst ammonia (*60*). Interestingly, the exchange times of the imino proton G22 extrapolates to relatively long base-pair lifetimes (5.2 ± 0.8 ms) for the C13/G22 Watson−Crick base pair in the limit of infinite exchange catalyst concentra-

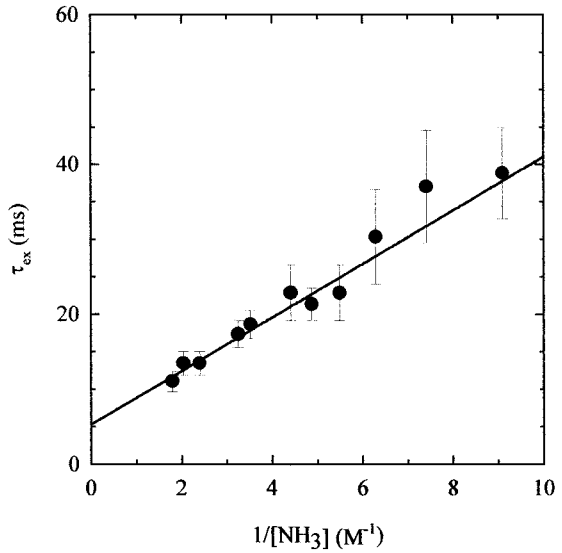


FIGURE 3: Exchange time τ_{ex} of the guanine imino proton G22 in the dA_5-bulge, plotted against the inverse ammonia base catalyst concentration at 15 °C. The regression line was calculated applying exchange times weighted according to their errors.

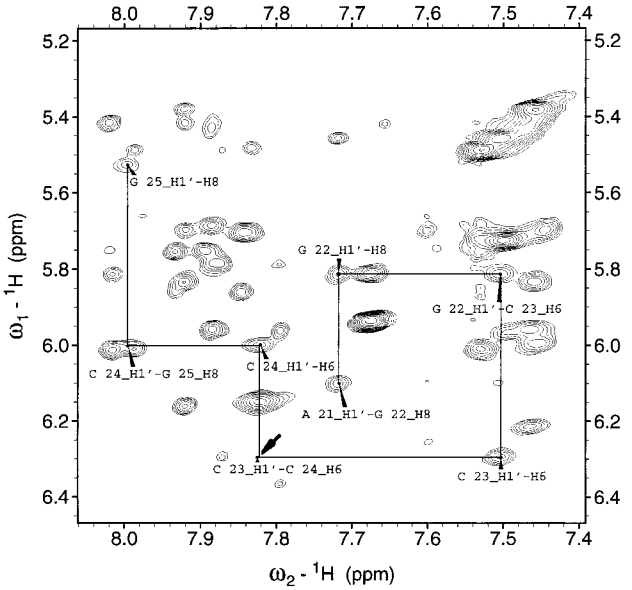


FIGURE 4: Expanded contour plot of the base to H1′ region of the NOESY experiment (200 ms mixing time) in D_2O buffer at 27 °C. The arrow indicates absent NOE for the C23C24 base step.

tion at 15 °C (Figure 3). The apparent dissociation constant αK_d is $(1.4 \pm 0.3) \times 10^{-6}$ for this base pair. The base-pair lifetime and apparent dissociation constant have values as expected for guanine imino protons that are unaffected by end-fraying and in B-like helix conformation (*60*).

(2) Nonexchangeable Proton Studies. The overall pattern of the NOE cross-peaks and the relative intensities are consistent with a right-handed double helix, a glycosidic torsion angle in the anti range, and S-type sugar pucker for both duplex parts (*58*). The well-resolved 2D NOE spectra allowed the sequential assignments of all H8, H6, H5, H2, methyl, H1′, H2′1, H2′2, and H3′ resonances via continuous base-to-sugar proton coupling networks at longer mixing times (cf. Supporting Information). Assignments of H4′ resonances are based on their strong intranucleotide cross-peak to H1′ and were further confirmed by scalar connectivities to H3′ in a 2QF-COSY spectrum.

The base H8/H6 to sugar H1′ cross-peaks of the NOESY spectrum of the dA_5-bulge indicate that all χ torsion angles are in the anti conformation because no strong intranucleotide H6/H8 to sugar H1′ NOE cross-peaks, characteristic of the syn conformation, are observed. Sequential NOE connectivities between aromatic H6/H8 protons and neighboring H1′, H2′1, and H2′2 sugar protons are observed continuously through the bulge loop between G6 and A10. The H8-H2′1 NOEs for base step A10-A11 and A11-G12 are very week or even absent. An interesting feature at base step A10-A11 is also the upfield shift (3.95 ppm) of the H4′ proton of A10 (cf. Supporting Information). Such an upfield shift has been observed for sugar protons of hairpin loops and has been related to the stacking of unpaired residue sugars over the next 3′ bases, producing an upfield ring-current shift of H4′ protons (*45*).

While the sequential H8 to sugar H1′ cross-peaks have normal intensities in the range of residue G6 to residue G12 in the bulge loop, the sequential cross-peak C24H6−C23H1′ on the opposite strand is not detectable even at longer mixing times (Figure 4). Furthermore, the sequential NOE cross-peaks C24H6−C23H2′1 and C24H6−C23H2′2 are either completely absent or very weak (data not shown), whereas the H6/H5 to H3′ sequential NOEs between C23 and C24 were still observed. The above data indicate that the sequential stacking is broken at base step C23C24. This break in sequential NOE connectivity between residues, which are structurally equivalent to C23 and C24, has been found before in smaller three-nucleotide DNA bulges (*59*).

Most of the adenine H2 resonances are easily identified by their comparatively long spin−lattice relaxation times with T_1 values between 3.4 and 4.4 s at 27 °C. They exhibit various weak intranucleotide and sequential contacts to H1′ sugar protons, particularly at longer mixing times. These include cross-peaks to their own H1′ and to H1′ of the 3′-linked nucleotide. For further confirmation of the proton assignment, we changed the adenine (A9) in the center of the bulge loop to a thymine. By introduction of the thymine a relatively easy assignment of the proton resonances in the bulge loop sequence was possible (Gollmick et al., unpublished results). The proton resonances of the bulge loop residues G6, A7, A10 (except H4′), A11, G12, C23, and C24 remain unchanged upon adenine−thymine substitution of residue 9.

Interproton distance restraints were established using nine MARDIGRAS distance sets (cf. Materials and Methods) which were calculated from NOE intensities of the three mixing times using three rotation correlation times and a starting structure derived from systematic conformational search in the program JUMNA, version 10.0 (cf. Materials and Methods). The 482 NMR-derived interproton distances are comprised of 333 intraresidue and 149 sequential distances (Table 1).

Determination of the Torsion Angles. The H1′−H2′1 cross-peaks in the 2QF-COSY spectrum indicate that all nonterminal residues are in the C2′-endo sugar conformation domain. The $^3J_{H1'-H2'1}$ coupling constants were found to vary between 8.8 and 11.7 Hz. Furthermore, the H2′2-H3′ cross-peaks are absent for all nonterminal residues. The deoxyribose pseudorotational phase angles were constrained to 90−190° and to 110−180° for all sugars showing $^3J_{H1'-H2'1}$

Table 1: Structural Statistics of the 15 NMR Conformers Derived from TAD Calculations Representing the Solution Structure of the DA_5-bulge DNA Molecule

NOE distance restraints	482
sequential distances	149
intraresidue distances	333
non-NOE distance restraints	13
hydrogen bonding restraints	26
torsion angle restraints	235
average restraints per residue	26
DYANA target function ($Å^2$)	1.49 ± 0.05
distance limit deviations > 0.2 Å	0
dihedral angle deviations > 5 °	0
mean pairwise RMSD[a],all heavy atoms (Å)	1.38 ± 0.71
RMSD[a] relative to mean structure, all heavy atoms (Å)	0.91 ± 0.55

[a] All nonterminal residues.

coupling constants of >8.8 and >10.5 Hz, respectively. The puckering amplitude was set to 36° for all calculations of the sugar torsion angles. For the terminal sugars, which presumably exhibit a more complex dynamic behavior caused by end fraying, no sugar torsion angle constraints were applied.

A set of 160 backbone torsion angles was derived from the qualitative and quantitative evaluation of the ^{31}P-^{1}H heteronuclear correlation spectrum and the 2D NOE spectra (*2*, *4*, *6*, *45*, *48*). We detected safely 27 $^{3}J_{H3'-P}$ and 25 $^{4}J_{H4'-P}$ cross-peaks in the ^{1}H-^{31}P correlation spectrum (cf. Supporting Information). Only the $^{4}J_{H4'-P}$ cross-peaks of residues C23 and C24 could not be identified. The ability to detect the long-range four-bond $^{4}J_{H4'-P}$ coupling indicates that the four consecutive bonds in the H4′−C4′−C5′−O5′−P backbone linkage lie in the same plane forming a W-shaped conformation. This is the case in B-form DNA where the β and γ torsion angles are in trans and gauche$^+$ domain, respectively (*48*). The trans conformation of the β torsion angle of the bulge loop residues was further verified by experimentally derived H5′1 and H5′2 line widths showing values between 21 and 25 Hz (*44*). The H4′ line widths varying between 8 and 14 Hz for all bulge loop residues (G6 to G12) suggest gauche$^+$ conformation of the γ torsion angle. On the basis of these observations, a β torsion angle range 120−240° and a γ torsion angle range 10−110° was used for all residues, except those of residue C23 and C24.

The observation of 27 $^{3}J_{H3'-P}$ cross-peaks, the absence of detectable four bond $^{4}J_{H2'1-P}$, and all sugars being predominantly C2′-endo enable us to conclude that all residues have ϵ values in the trans conformation (150−210°), except those for residues G1, G17, C18, and C29 (*45*, *49*). The trans conformation of the ϵ torsion angle of the bulge loop residues was further verified by experimentally derived sums of J couplings on H3′ showing values between 9.3 and 10.7 Hz (*44*).

All ^{31}P signals of the residues resonate in the range between −0.15 and −0.95 ppm, indicating the absence of unusual phosphate dihedral angles. Thus, for the residues in the duplex parts, the range of α and ζ torsion angles was set to $-60 \pm 30°$ and $-90 \pm 60°$, respectively (*1*, *50*, *61*). Additionally, the α and ζ torsion angles in the bulge loop from residue G6 to residue A11 were restrained to $0 \pm 120°$. Furthermore, the α and ζ torsion angle between residue C23 and C24 were also restrained to $0 \pm 120°$.

The final restraint data set (Table 1) included 482 distance restraints, 13 non-NOE distance restraints, 26 hydrogen bonding restraints, and 235 torsion angle restraints; the total number 756 corresponds to an average of 26 restraints per nucleotide residue.

Structure Determination. On the basis of the experimentally derived restraint data set the three-dimensional structure of the dA_5-bulge was calculated using the software package DYANA. The DYANA target function values of the best 15 conformers (shown in Figure 5) are in the range 1.41−1.56 $Å^2$. This indicates the consistency of the input data and the low deviations from the constraints in the resulting conformers. The RMSD for all heavy atoms relative to the mean structure of the final 15 energy minimized conformers is 0.91 ± 0.55 Å for all nonterminal residues (Table 1).

(1) Base−Base Stacking in the Bulge Loop. The NMR-derived solution structure of the dA_5-bulge shows that all bulged adenine residues are localized at intrahelical positions within the double helical stems (Figure 5). This concurs with NMR structures of three-nucleotide DNA bulges (*28*, *59*), showing that the unpaired bases stack inside the helix. The first 5′-adenine in the bulge loop stacks well on the end of the double helical stem. The stacking of the bulge residues A7, A8, A9, and A10 is concluded from the presence of all H8/H6 to H1′, H2′1, and H2′2 NOE connectivities in this bulge loop region. A kink in the backbone of the dA_5-bulge is observed between A10 and A11, expressed by a ζ torsion angle in the *gauche*$^+$ range for A11. The local conformation of the base step A10A11 is characterized by the stacking of the A10 sugar over the A11 base; this is concluded from the upfield shift of the H4′ proton of A10. In an earlier study (*13*), the authors studied the solution structure of dA_5-bulge loops using small molecular probes that specifically react with single-stranded nucleotides. In accordance with our results, the adenine at the 5′-end (A7) of the bulge stacks well on the end of the duplex. In addition, and again in agreement with their data, A10 and A11 are solvent exposed due to the disruption of stacking between these two bases in the bulge loop.

Across from the bulge loop site, loss of base stacking is observed at base step C23C24. The sequential NOEs H6 to H1′ and H6 to H2′1/H2′2 between residues C23 and C24 were absent whereas the H6/H5 to H3′ sequential NOEs were still observed resulting in a spatial shearing apart of these bases and a loss of stacking interaction. This was also observed for the three-nucleotide DNA bulges (*28*, *59*).

An extraordinary hydrogen bond is observed in about 60% of the conformers (Figure 6). The amino proton of C24, which is not involved in base pairing, is in close contact with one of the phosphate oxygen atoms of C23. This observation is in agreement with the 2D NOE spectrum in H_2O. The aforementioned exchangeable proton of C24 shows the highest downfield shift of all comparable protons (cf. Supporting Information) and can therefore be involved in hydrogen bonding. The formation of this hydrogen bond is consistent with the sequence dependence of single base bulge gel mobility experiments (*15*). It was observed that single bulged DNA molecules with dG/dC neighboring base pairs migrate slower in gels than all other analyzed sequence combinations. A reduced mobility in polyacrylamide gels is interpreted in terms of local kinks at the bulge sites. Only in the case of a 5′ framing dG/dC base pair a hydrogen bond between the phosphate backbone and one of the flanking base pairs can be formed. Thus, these sequence constructions

FIGURE 5: Stereoview of the 15 TAD-derived NMR conformers of the dA_5-bulge. Terminal base pairs and all hydrogen atoms are omitted for clarity. The 15 NMR conformers were deposited at the Brookhaven Protein Data Bank (ID code: 1QSK).

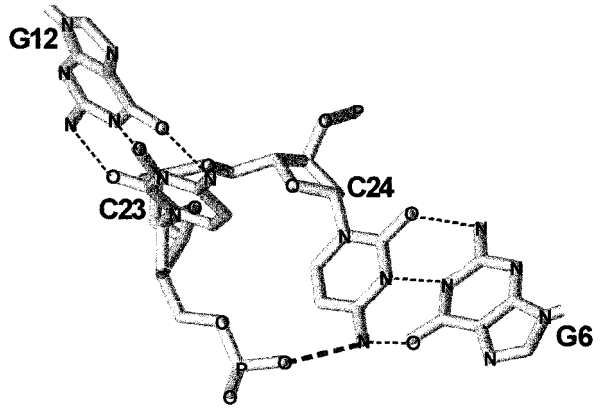


FIGURE 6: Relative position of the two base pairs flanking the five unopposed adenines. The unusual hydrogen bond between the phosphate oxygen (OP2) of base C24 and the amino proton of C23 is shown as bold dotted line.

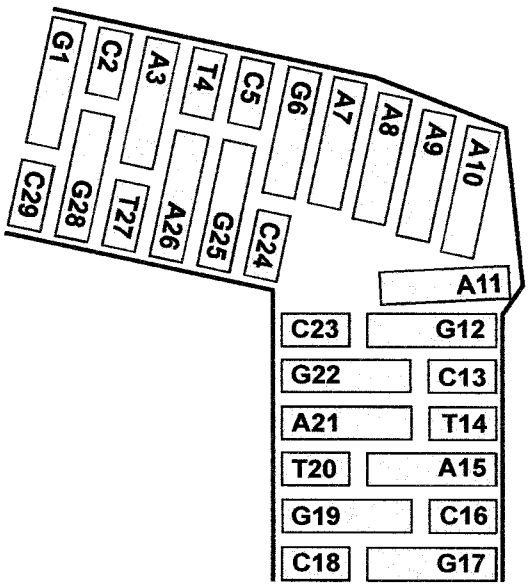


FIGURE 7: Schematic illustration of the dA_5-bulge showing the base stacking in the bulge loop region.

are assumed to be either more kinked or more stable than others, experiencing a larger friction during gel electrophoresis.

The dA_5-bulge structure can be understood in the following simple terms (Figure 7): The unpaired purine bases stack on the framing stems and on themselves in rough continuation of the double helical axis. Due to stereochemical constraints on the opposite strand, the stacking must be disrupted. This disruption is observed after four 5′-loop adenines (between A10 and A11), defining a kinking angle between these two bases. The total rise of these four stacked loop adenines equals roughly the diameter of the double helix. The fifth loop adenine A11 partially stacks on the other stem. The kinking angle between A10 and A11 reflects the kinking angle of the two stems.

(2) Kinking Angle and Helical Parameters. The NMR-derived solution structure of the dA_5-bulge DNA molecule shows an overall kinking angle of 73° (±11°) that is in agreement with other experimental data for five-nucleotide bulge loops. Previous studies (*13, 14*) of DNAs containing bulged bases on one of the two strands of a duplex showed that they exhibit slower electrophoretic mobility than non-bulged DNAs, indicating that bulges induce kinks into the DNA. Fluorescence energy transfer measurements (FRET) proposed a pronounced kink in the DNA helix of 85−105° for a five-adenine DNA bulge (*18*). Molecular modeling studies (*40*) predicted a kinking angle of 87° (±8°) for the dA_5-bulge. Furthermore, for a five-adenine RNA bulge, a kinking angle of 75° (±3°) was deduced from transient electric birefringence measurements (*16*). The recent NMR-derived solution structure of an RNA bulge with five unpaired bases AAUAA (*29*) shows a kinking angle of 90° (±14°). The kinking angle of the NMR-based solution structure of the dA_5-bulge is slightly lower (73 ± 11°) compared to this five-nucleotide RNA-bulge.

The helical parameters describing the mutual orientation of the two bulge framing base pairs (and thus the two helices) are *shift* −8.2 Å (±0.7 Å), *slide* 0.3 Å (±0.7 Å), *rise* 12.7 Å (±0.8 Å), *tilt* 68.1° (±4.2 °), *roll* 29.1° (±7.6°), and *twist* 105.2° (± 6.6°).

(3) Flexibility in the Bulge Loop Region. Our NMR data indicate clearly that the bulge loop region forms a well-defined structure. Every sugar in the bulge loop region has a $^3J_{H1'-H2'1}$ coupling between 9 and 12 Hz, which is consistent with a major C2′-endo conformation. No significant contribution of C3′-endo sugar puckers is observed. The same range of $^3J_{H1'-H2'1}$ coupling constants was also found for the nonterminal residues in the stem regions of the dA_5-bulge.

For evaluation of the local base dynamics in the bulge loop region we examined the temperature-dependent line width of the adenine H2 resonances using a recovery pulse sequence as previously described (*37*). Only the H2 resonance line of residue A11 was noticeably broadened (by a factor of 2−4) between 20 and 35 °C in comparison to all other adenine H2 in the stems and bulge loop region of the dA_5-bulge. This line broadening is compatible with a partial motion of the base A11 as expected by the interruption of the stacking interaction between A10 and A11.

(4) Comparison of the NMR Structure with FRET Experiments of Double-Bulge DNA Molecules. It has been shown by FRET measurements and molecular modeling (*30*, *40*) that larger DNA molecules containing two dA_5-bulges occur in a nearly planar U-shaped conformation if nine base pairs separate these bulges. Molecular modeling studies of these DNA constructs show that these observations are in accord with the NMR structure. By superimposing stem 1 with stem 2 of two identical dA_5-bulge NMR conformations on the computer screen, a double-bulge DNA is generated which exhibits an approximately planar conformation when 8−9 base pairs form the middle stem. The kinking angle of 84° (±3°) for the single dA_5-bulge in such a double-bulge DNA construct obtained by FRET measurements (*30*) agrees well with the kinking angle observed for the NMR structure (73 ± 11°). Thus, in case of the dA_5-bulge local (up to 5 Å) and global (10−100 Å) distance restraints obtained with two complementary biophysical methods (NMR and FRET, respectively) are in concord with each other.

ACKNOWLEDGMENT

The authors thank E. Birch-Hirschfeld for the synthesis of the DNA oligonucleotides, W. Scheiding and J. Flemming for skillful technical assistance, and M. Leijon for help with measurements of the base-pair dynamics. We are grateful to S. Wartewig for using his Raman equipment. This study was supported by fond of the Chemical Industry (H.F.), the Deutsche Forschungsgemeinschaft (DFG), the Deutsche Akademische Austauschdienst (DAAD), and the Thuringian Ministry of Science, Research, and Culture (TMWFK).

SUPPORTING INFORMATION AVAILABLE

One table giving ^1H and ^{31}P chemical shifts. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Streisinger, G., Okada, Y., Emrich, Y., Newton, J., Tsugita, A., Terzaghi, E., and Inouye, M. (1966) *Cold Spring Harbor Symp. Quant. Biol. 31*, 77−84.
2. Nelson, J. W., and Tinoco, I. (1985) *Biochemistry 24*, 6416−6421.
3. White, S. A., and Draper, D. E. (1987) *Nucleic Acids Res. 15*, 4049−4064.
4. Stassinopoulos, A., Ji, J., Gao, X., and Goldberg, I. H. (1996) *Science 272*, 1943−1946.
5. Flor, P. J., Flanegan, J. B., and Cech, T. R. (1989) *EMBO J. 11*, 3391−3399.
6. Cate, J. H., Gooding, A. R., Podell, E., Zhou, K., Golden, B. L., Kundrot, C. E., Cech, T. R., and Doudna, J. A. (1996) *Science 273*, 1678−1685.
7. Peattie, D. A., Douthwaite, S., Garrett, R. A., and Noller, H. F. (1981) *Proc. Natl. Acad. Sci. U.S.A. 78*, 7331−7335.
8. Romaniuk, P. J., Lowary, P., Wu, H.-N., Stormo, G., and Uhlenbeck, O. C. (1987) *Biochemistry 26*, 1563−1568.
9. Wu, H.-N., and Uhlenbeck, O. C. (1987) *Biochemistry 26*, 8221−8227.
10. Weeks, K. M., and Crothers, D. M. (1991) *Cell 66*, 577−588.
11. Puglisi, J. D., Tan, R., Calnan, B. J., Frankel, A. D., and Williamson, J. R. (1992) *Science 257*, 76−80.
12. Aboul-ela, F., Karn, J., and Varani, G. (1995) *J. Mol. Biol. 253*, 313−332.
13. Bhattacharyya, A., and Lilley, D. M. J. (1989) *Nucleic Acids Res. 17*, 6821−6841.
14. Hsieh, C.-H., and Griffith, J. D. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 4833−4837.
15. Wang, Y.-H., and Griffith, J. D. (1991) *Biochemistry 30*, 1358−1363.
16. Zacharias, M., and Hagerman, P. J. (1995) *J. Mol. Biol. 247*, 486−500.
17. Luebke, K. J., and Tinoco, I. (1996) *Biochemistry 35*, 11677−11684.
18. Gohlke, C., Murchie, A. I. H., Lilley, D. M. J., and Clegg, R. M. (1994) *Proc. Natl. Acad. Sci. U.S.A. 91*, 11660−11664.
19. Lilley, D. M. J. (1995) *Proc. Natl. Acad. Sci. U.S.A. 92*, 7140−7142.
20. Wang, Y.-H., Barker, P., and Griffith, J. D. (1992) *J. Biol. Chem. 267*, 4911−4915.
21. VandenHoogen, Y. T., Van Beuzekom, A. A., de Vroom, E., Van der Marel, G. A., Van Boom, J. H., and Altona, C. (1988) *Nucleic Acids Res. 16*, 5013−5031.
22. Morden, K. M., Chu, Y. G., Martin, F. H., and Tinoco, I. (1983) *Biochemistry 22*, 5557−5563.
23. Morden, K. M., Gunn, B. M., and Maskos, K. (1990) *Biochemistry 29*, 8835−8845.
24. Joshua-Tor, L., Rabinovich, D., Hope, H., Frolow, F., Appella, E., and Sussmann, J. L. (1988) *Nature 334*, 82−84.
25. Nikonowicz, E. P., Meadows, R. P., and Gorenstein, D. G. (1990) *Biochemistry 29*, 4193−4204.
26. Woodson, S. A., and Crothers, D. M. (1989) *Biopolymers 28*, 1149−1177.
27. Rosen, M. A., Live, D., and Patel, D. J. (1992) *Biochemistry 31*, 4004−4014.
28. Aboul-ela, F., Murchie, A. I. H., Homans, S. W., and Lilley, D. M. J. (1993) *J. Mol. Biol. 229*, 173−188.
29. Luebke, K. J., Landry, S. M., and Tinoco, I. (1997) *Biochemistry 36*, 10246−10255.
30. Stühmeier, F., Hillisch, A., Diekmann, S., and Clegg, R. M. (1999) in *Protein-DNA Interactions: A Practical Approach* (Travers, A. A., and Buckle, M., Eds.) IRL Press at Oxford University Press, Oxford (in press).
31. Birch-Hirschfeld, E., Földes-Papp, Z., Gührs, K. H., and Seliger, H. (1996) *Helv. Chim. Acta 79*, 137−150.
32. Dornberger, U., Behlke, J., Birch-Hirschfeld, E., and Fritzsche, H. (1997) *Nucleic Acids Res. 25*, 822−829.
33. Hore, P. J. (1983) *J. Magn. Reson. 55*, 283−300.
34. Rance, M., Sorensen, O. W., Bodenhausen, G., Wagner, G., Ernst, R. R., and Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun. 117*, 479−485.
35. Sklenar, V., Miyashiro, H., Zon, G., Miles, T., and Bax, A. (1986) *FEBS 208*, 94−98.
36. Suzuki, E. I., Pattabiraman, N., Zon, G., and James, T. L. (1986) *Biochemistry 25*, 6854−6865.
37. McAteer, K., Ellis, P. D., and Kennedy, M. A. (1995) *Nucleic Acids Res. 23*, 3962−3966.
38. Gonzáles, C., Stec, W., Kobylanska, A., Hogrefe, R. I., Reynolds, M., and James, T. L. (1994) *Biochemistry 33*, 11062−11072.
39. Borgias, B. A., and James, T. L. (1990) *J. Magn. Reson. 87*, 475−487.
40. Hillisch, A. (1998) Thesis, University of Vienna.
41. Lavery, R. (1988) in *DNA Bending & Curvature* (Olson, M. H., Sarma, M. H., Sarma, R. H., and Sundaralingam, M., Eds.) pp 191−211, Adenine Press.
42. Weiner, S. J., Kollmann, P. A., Nguyen, D. T., and Case, D. A. (1986) *J. Comput. Chem 7*, 230−252.

43. Weisz, K., Shafer, R. H., Egan, W., and James, T. L. (1994) *Biochemistry 33*, 354−366.
44. Kim, S.-G., Lin, L.-J., and Reid, B. R. (1992) *Biochemistry 31*, 3564−3574.
45. Chou, S.-H., Zhu, L., Gao, Z., Cheng, J.-W., and Reid, B. R. (1996) *J. Mol. Biol. 264*, 981−1001.
46. Wijmenga, S. S., Mooren, M. W., and Hilbers, C. W. (1995) in *NMR of Macromolecules: A Practical Approach* (Roberts, G. C. K., Ed.) pp 217−357, IRL Press at Oxford University Press, Oxford.
47. Collini, M., Chirico, G., Baldini, G., and Bianchi, M. E. (1998) *Macromolecules 31*, 695−702.
48. Altona, C. (1982) *Recl. Trav. Chim. Pays-Bas 101*, 413−433.
49. Blommers, M. J. J., Van de Ven, F. J. M., Van der Marel, G. A., Van Boom, J. H., and Hilbers, C. W. (1991) *Eur. J. Biochem. 201*, 33−51.
50. Gorenstein, D. G., Schroeder, S. A., Fu, J. M., Metz, J. T., Roongta, V., and Jones, C. R. (1988) *Biochemistry 27*, 7223−7237.
51. Güntert, P., and Wüthrich, K. (1991) *J. Biomol. NMR. 1*, 447−456.
52. Güntert, P., Mumenthaler, C., and Wüthrich, K. (1997) *J. Mol. Biol. 273*, 283−298.
53. Pearlman, D. A., Case, D. A., Caldwell, J. W., Ross, W. S., and Cheatham, T. E., III (1995) AMBER 4.1. University of San Francisco.
54. Koradi, R., Billeter, M., and Wüthrich, K. (1996) *J. Mol. Graphics 14*, 51−55.
55. Lavery, R., and Sklenar, H. (1989) *J. Biomol. Struct. Dyn. 6*, 655−667.
56. Taillandier, E., and Liquier, J. (1992) *Methods Enzymol. 211*, 307−335.
57. Peticolas, W. L. (1995) *Methods Enzymol. 246*, 389−416.
58. Wüthrich, K. (1986) *NMR of proteins and nucleic acids*, John Wiley & Sons, Inc., New York.
59. Rosen, M. A., Shapiro, L., and Patel, D. J. (1992) *Biochemistry 31*, 4015−4026.
60. Gueron, M., and Leroy, J.-L. (1995) *Methods Enzymol.* 383−413.
61. Saenger (1984) *Principles of nucleic acid structure*, Springer-Verlag, New York.

BI9906874