A Pseudosquare Knot Structure of DNA in Solution†

Nikolai B. Ulyanov, Valery I. Ivanov,‡ Elvira E. Minyat,‡ Elena B. Khomyakova,‡ Maria V. Petrova,‡ Krystyna Lesiak,§ and Thomas L. James*

Department of Pharmaceutical Chemistry, University of California, San Francisco, California 94143-0446, Engelhardt Institute of Molecular Biology, Russian Academy of Sciences, Moscow, Russia, and Food and Drug Administration, Center for Biologics Evaluation and Research, Division of Allergenic Products and Parasitology, Laboratory of Biophysics, 1401 Rockville Pike, Rockville, Maryland 20852

Received May 4, 1998; Revised Manuscript Received July 10, 1998

ABSTRACT: We report a high-resolution NMR structure of a homodimer formed by a synthetic 25 residue DNA oligonucleotide GCTCCCATGGTTTTTGTGCACGAGC. This structure presents a novel structural motif for single-stranded nucleic acids, called a pseudosquare knot (PSQ). The oligonucleotide was originally designed to mimic a slipped-loop structure (SLS), another "unusual" DNA structure postulated as an alternative conformation for short direct repeats in double-stranded DNA. The design of the sequence is compatible with both SLS and PSQ structures, both of which possess identical sets of base-paired and unpaired nucleotides but different tertiary folds. We used deuteration of the H8 positions of purines to ascertain that the PSQ is actually formed under the conditions used. The PSQ structure was solved based on homonuclear proton nuclear Overhauser effect data using complete relaxation matrix methods. The structure essentially consists of two side-by-side helices connected by single-stranded loops. Each of the helices is well-defined; however, the relative orientation of the two remains undetermined by the NMR data. The sequences compatible with the PSQ formation are frequent in single-stranded genomes; this structure may play a role as a dimerization motif.

Single-stranded nucleic acids can form a variety of conformations with specific shapes and chemical properties, depending on their sequence. Here, we describe a novel tertiary fold, a pseudosquare knot (PSQ),[1] for nucleic acids with certain sequence motifs, and we report a high-resolution NMR structure of a DNA 25-mer adopting this conformation in solution. Our interest in this structure has originated from the studies of so-called unusual DNA conformations; however, this novel type of folding may have a significance for single-stranded genomes in general, both DNA and RNA.

The term "unusual DNA structures" has been traditionally used for noncanonical, alternatives to B- or A-form conformations, which can be adopted by certain DNA sequences under certain conditions (*1, 2*). Some well-known examples of such conformations are left-handed Z-DNA (*3*), H-DNA triplex (*4*), cruciform, or Holliday intermediate (*5*). Each of these structures requires a defined nucleotide sequence motif; for example, a palindrome, or inverted repeat, is necessary for the cruciform formation. Short direct repeats constitute another sequence motif frequently found in genomes, including regulatory regions. An alternative structure, SLS, had been proposed for this motif, in an attempt to explain the data on DNA cleavage by a single strand-specific nuclease S1 (*6, 7*). SLS is thought to be formed as a result of strand slippage involving two repeats, which leads to formation of two shifted loops protruding from the opposite strands. If the two repeats were separated by about a half-turn of the double helix, then the loops are positioned on the same side of DNA duplex, and the loop nucleotides can potentially form tertiary base pairs (*7, 8*). This circumstance allows one to classify SLS as a special type of pseudoknot, even though it is very different from the classical coaxial pseudoknots (*9*).

Unusual DNA structures, including SLS, often include a combination of base-paired and unpaired nucleotides. Because of that, such structures are typically high-energy conformations relative to duplexes; to be formed, they must be stabilized by appropriate conditions. For example, alternating GC sequences can be converted into left-handed Z-form by high ionic strength, and cruciforms can be extruded from circular DNA by increased superhelicity. Dependence on superhelicity is a common feature of many unusual DNA structures, which implicates all of them in the regulation of gene expression (*1, 10*). At the same time, this makes such conformations a difficult object for high-resolution structural studies, because they need to compete with low-energy fully base-paired duplexes. An alternative useful approach to study them is to design a sequence

† This work was supported financially by the National Institutes of Health Grants GM 39247 and RR 01081, Fogarty International Research Collaboration Award TW 00438, Russian Fund for Basic Research Grant 95-0411707, and Russian State Support of Scientific Schools Grant 980903.

* To whom correspondence should be addressed at UCSF. Fax: (415) 502-4690. E-mail: james@picasso.ucsf.edu.

‡ Russian Academy of Sciences.

§ Center for Biologics Evaluation and Research. Present address: Codon Pharmaceuticals, Inc., 200 Perry Parkway, Gaithersburg, Maryland 20877.

[1] Abbreviations: SLS, slipped-loop structure; PSQ, pseudosquare knot; NMR, nuclear magnetic resonance; NOE, nuclear Overhauser effect; NOESY, nuclear Overhauser effect spectroscopy; 1D, one-dimensional; 2D, two-dimensional; DNA, deoxyribonucleic acid; RNA, ribonucleic acid.

S0006-2960(98)01018-6 CCC: $15.00
Published on Web 08/22/1998

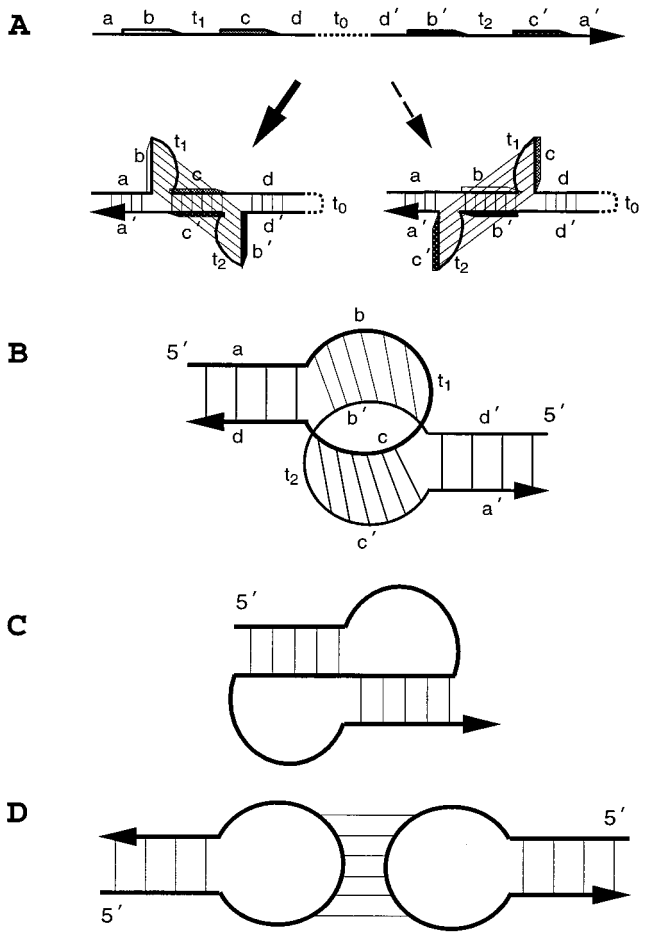

FIGURE 1: Schematic representation of secondary structures of SLS-related conformations. Thin lines show base pairs. (A) Design of oligonucleotide that forms immobile SLS. Segments with complementary sequences are denoted by apostrophes. Segments $\mathbf{t_1}$, $\mathbf{t_2}$, and $\mathbf{t_0}$ remain unpaired in SLS. Minihelices **b:b′** and **c:c′** correspond to short direct repeats in the case of SLS formation for double-stranded DNA. However, **b** and **c** must have a different sequence in order to make SLS immobile. Two opposite slippage directions could potentially lead to two SLS conformers with different pattern of stacking interaction; however, formation of tertiary base pairs is feasible for one SLS conformer only, shown in left panel (see text). Only this conformer is shown for all SLS isomers in Figure 2. (B) Pseudosquare knot (PSQ) or "kissing SLS". Note the similarity with the square (or reef) knot. The SLS-type arrangement of helical segments is apparent from comparison with Figure 2. Notations for the segments **a**, **b**, **c**, etc., are the same as in panel A. This type of folding is not possible for direct repeats in double-stranded DNA or for single-stranded sequences with central hairpin loop $\mathbf{t_0}$ (panel A). (C) Classical pseudoknot. (D) Kissing hairpins. In contrast to PSQ, there is only one continuous stretch of base pairs in the loop regions.

forming an immobile unusual conformation. For example, construction of immobile Holliday junctions allowed a number of physical-chemical and structural studies on this structure (*11*−*13*).

An immobile SLS construct (SLS palindrome) is schematically shown in Figure 1A; such sequences can potentially form SLS but not the plain duplex (*8*, *14*, *15*). The core of this construct is composed of fragments **b**-$\mathbf{t_1}$-**c** and **b′**-$\mathbf{t_2}$-**c′**, which form two mini-helices, **b:b′** and **c:c′** (five or six base pairs each), and two loops, $\mathbf{t_1}$ and $\mathbf{t_2}$. Terminal stems **a:a′** and **d:d′** as well as the optional hairpin loop $\mathbf{t_0}$ serve to stabilize the core. Fragments **b** and **c** correspond to short direct repeats in the case of SLS formed by strand slippage in a DNA duplex. To make SLS immobile, **b** and **c** must have different sequence; also the single-stranded segments $\mathbf{t_1}$ and $\mathbf{t_2}$ must be noncomplementary to each other. This prevents base pairing of strands **b** and **c′**, $\mathbf{t_1}$ and $\mathbf{t_2}$, and **c** and **b′**, and consequently, it makes formation of a plain stem-and-loop structure impossible. There are two potential slippage directions (both for short direct repeats in duplexes and for immobile SLS palindromes). Model-building studies demonstrated that only one slippage direction is possible for both DNA and RNA (*14*, *16*) (Figure 1A, left panel). This prediction has been shown to be true in our present work for a model DNA oligonucleotide. Chemical and S1 nuclease probing (*8*, *15*) of various synthetic immobile SLS palindromes showed that their secondary structures are indeed consistent with the SLS conformation.

Our first low-resolution NMR studies of SLS (*14*) have been carried out for a model 55 residue DNA oligonucleotide, which had a hairpin loop $\mathbf{t_0}$ (Figure 1A). The oligonucleotide was designed with a GC-rich sequence in such a way that AT pairs could be formed only via tertiary base pairs between segments **b** and **b′**; both **b** and **b′** had the self-complementary sequence CCATGG. Poor chemical shift dispersion prevented us from carrying out high-resolution NMR studies of that molecule. Nevertheless, we assigned a resonance at ca. 13.7 ppm to an AT pair imino proton, based on its chemical shift, and thus obtained critical evidence for SLS formation (*14*). To simplify proton NMR spectra, we altered our earlier design of the SLS palindrome for the present work. The central hairpin loop $\mathbf{t_0}$ (Figure 1A) was eliminated, so that the molecule was cut in two pieces, and the SLS formed as a dimer of two strands. Depending on sequence design, the two strands can be forced to fold as one of two possible conformations. If segment **a** is complementary to **a′** and **d** is complementary to **d′**, the two strands must fold as SLS proper, or "linear SLS", similar to that shown in Figure 1A. If segment **a** is complementary to **d** and **a′** complementary to **d′**, the strands must fold back and form a structure which we call a "pseudosquare knot" (PSQ) or "kissing SLS" (Figure 1B), using an analogy to a kissing hairpin (*17*). Secondary structures of the classical pseudoknot and kissing hairpin are shown schematically in Figure 1, panels C and D, respectively, to illustrate the analogy. For the present work, both strands had the identical sequence SLS-25 (Figure 2, top left). Segments **b** (CCATGG) and **c** (GTGCAC) were made self-complementary, and segments **a** and **d′** were made identical (GCTC) and complementary to **a′** and **d** (GAGC). With such a design, the structure is expected to be symmetric, so there are half as many NMR signals to resolve in a spectrum. An interesting consequence of the design is that SLS-25 can potentially fold as either one of two isomers, SLS or PSQ (Figure 2). Despite the seemingly different secondary structures, local conformations and arrangements of helical segments (with the exception of single-stranded pentathymidine loops $\mathbf{t_1}$ and $\mathbf{t_2}$) may be identical in the two isomers (Figure 2, right). All base pairs are spectroscopically equivalent in the two folding isomers (when using homonuclear NMR). For example, helical stem **a:a′** (G1C2T3C4:G22′A23′G24′C25′) in the linear SLS isomer is equivalent to **a:d** (G1C2T3C4:G22A23G24C25) in PSQ. To distinguish between the two folding possibilities, we carried out a deuteration study on a mutant of SLS-25; this study

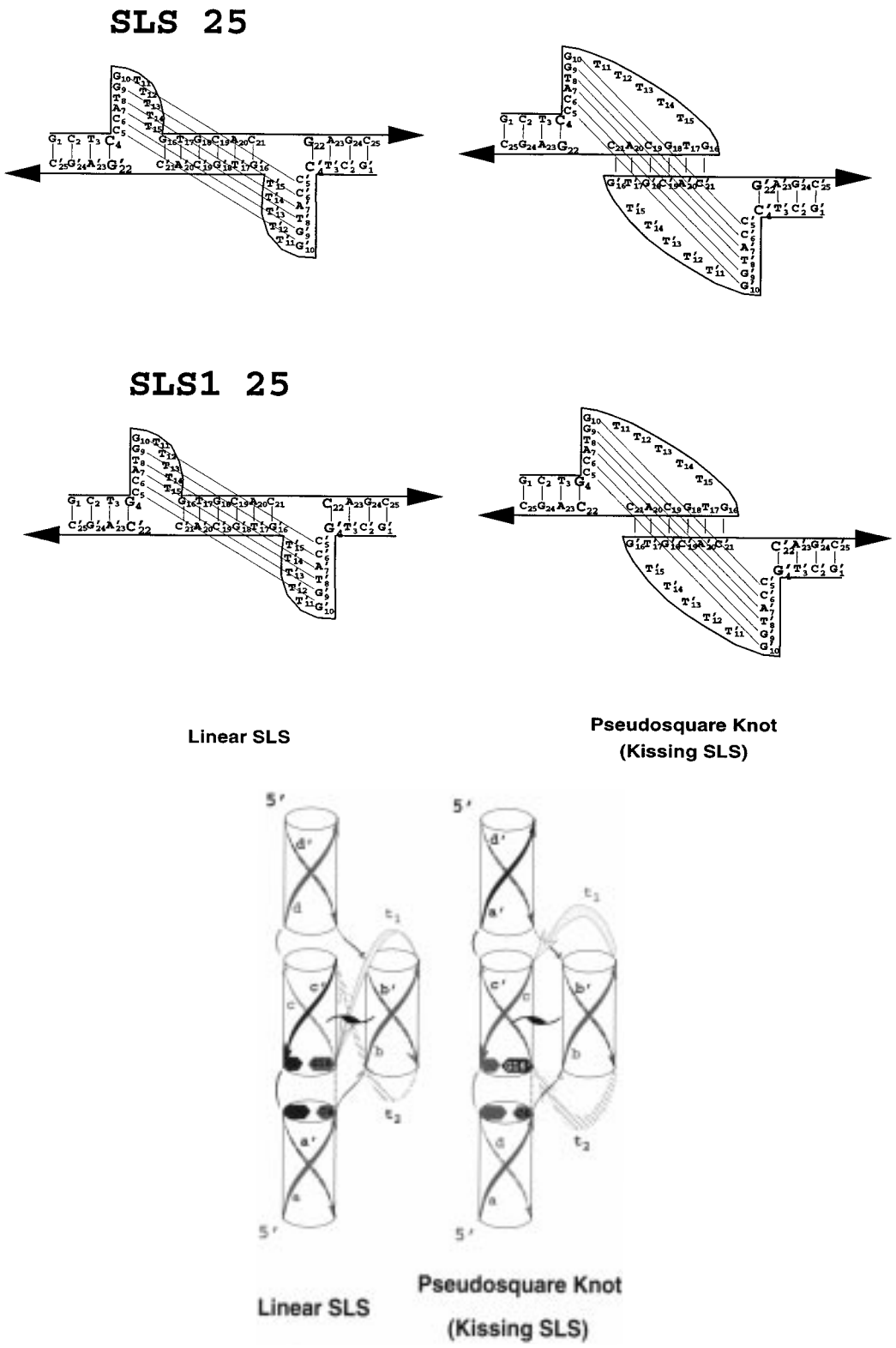

FIGURE 2: (Top left) Sequence and numbering scheme for a dimer formed by DNA oligonucleotide SLS-25, used for high-resolution NMR studies. Two potential folding isomers (SLS, left, and PSQ, right) with identical sets of base-paired and unpaired regions are shown. The conformer is shown with the stacking interaction corresponding to the left direction of strand slippage (Figure 1A). (Lower left) SLS (left) and PSQ (right) structures formed by SLS1−25, a mutated sequence used for the deuteration studies. Compared to SLS-25, SLS1−25 has a base pair G4:C22 instead of C4:G22 (shown in bigger font). (Right) Spatial arrangement of sequence segments in the two folding isomers of SLS, linear SLS (left) and PSQ (right). Double-helical segments are shown as cylinders; notations for the segments **a**, **b**, **c**, etc., are the same as in Figure 1B. One strand is colored green and yellow, the other strand is colored blue and magenta. Note that both isomers are symmetric and can have identical local conformations for all segments, except loops $\mathbf{t_1}$ (yellow) and $\mathbf{t_2}$ (magenta). Both isomers have residues C4 and G16 stacked, shown by dashed lines.

demonstrated that the PSQ fold is adopted in solution under the conditions used. The high-resolution structure of the PSQ SLS-25 in solution is determined with homonuclear ^{1}H NMR and reported here. Our prior low resolution studies on

various SLS constructs and preliminary NMR results are reviewed elsewhere (*18*).

MATERIAL AND METHODS

DNA Synthesis. DNA oligonucleotide SLS-25 (Figure 2, top left) was synthesized on an Applied Biosystems model 380B DNA synthesizer according to the manufacturer's standard procedures. Details of the purification procedures are described elsewhere (*18*). Oligonucleotide SLS1−25 (Figure 2, lower left) was purchased from Integrated DNA Technologies.

Chemical Modification. The modification of oligonucleotides with potassium permanganate was performed at 4 °C in the presence of 10 mM $MgCl_2$ as described before (*8*, *15*); concentration of DNA was 0.04 mM in strands.

NMR Spectroscopy. NMR samples (SLS-25 or SLS1−25) were prepared by dissolving the lyophilized powder in a 5 mm "Shigemi" tube containing 0.27 mL of a buffer with 2.7 mM $MgCl_2$, 60 mM NaCl, 13 mM Na_2PO_4, and either D_2O or 90/10% mixture of H_2O and D_2O. D_2O samples were additionally lyophilized several times from increasingly higher grades of D_2O. DNA concentration was approximately 2 mM in strands; pH was adjusted with NaOH to 7.0 for D_2O samples and to 6.0 for water samples. All NMR experiments were run on either a Varian UnityPlus 600 MHz or a GE Omega 500 MHz spectrometer. 2D NOESY experiments (*19*) were performed at 10 °C using a hypercomplex phase cycling (*20*). A 2D NOESY spectrum was acquired for the H_2O SLS-25 sample at mixing time 300 ms using program SS-NOESY with a sine-shaped pulse for water suppression (*21*) with the excitation maximum in the middle of imino proton resonances of base-paired nucleotides. NOESY spectra for the D_2O SLS-25 sample were acquired at mixing times 75 and 150 ms. NOESY spectra for fully protonated and protonated/deuterated SLS1−25 in D_2O were acquired at mixing time 300 ms with presaturation of residual HDO signal. Spectra were processed with the program Striker (*22*) and assigned with the help of the program Sparky (*23*) run on either Sun Sparc2 or Silicon Graphics INDY R5000 work stations; NOE cross-peaks were integrated using Gaussian linefit and deconvolution of moderately overlapped peaks with Sparky.

Deuteration. H8 positions of purines in SLS1−25 were exchanged for deuterium by incubating the oligonucleotide in D_2O at 51 °C for 189 h; the exchange was at least 90% complete as monitored by ^1H NMR. The solvent for deuteration was a D_2O solvent similar to that used for NMR experiments, with the exception of $MgCl_2$, which was absent; pH was 8.3.

Structure Calculation. The NMR structures for SLS-25 were refined with the programs DYANA (*24*), AMBER (*25*), and DNAminiCarlo (*26*). The initial model for SLS-25 was built with DNAminiCarlo as described before (*14*). The DNAminiCarlo program uses the DNA helical parameters as independent variables; it first places the bases in the specified positions, and then calculates the sugar−phosphate backbone by a chain-closure algorithm (*27*). This approach allows convenient manipulations with helical segments of the molecule while maintaining perfect stereochemistry.

The initial PSQ model and integrated NOE intensities from the two NOESY datasets in D_2O (excluding intensities involving H4′, H5′ or H5″ protons) were used as input for the MARDIGRAS program (*28*). This program determines interproton distances from NOE data accounting for a complete network of relaxation pathways in a molecule. Distance restraints involving nonexchangeable protons were calculated by MARDIGRAS with the random error analysis procedure (*29*). At first, MARDIGRAS was run for a series of correlation time values, and then a correlation time was selected which best reproduced the fixed interproton distances and certain distances with small allowed variation (mainly intrasugar distances). Distance bounds calculated with the "random error MARDIGRAS" program were used to refine a preliminary structure with restrained Monte Carlo and minimization routines of DNAminiCarlo (*30*). This preliminary structure was then used in the second, final, round of MARDIGRAS analysis. For each of the D_2O NOESY datasets at mixing time of 75 and 150 ms and an overall correlation time 15 ns, MARDIGRAS was run 50 times with input NOE intensities randomly modified to account for the experimental noise (assumed equal to the smallest integrated intensity in the dataset) and integration errors (10% on average). From the resulting distributions for each interproton distance, 10% of smallest and 10% of longest values were discarded; the range of the remaining values determined the lower and upper bounds for each distance. A total of 656 distance restraints (per dimer of SLS-25) have been determined for nonexchangeable protons; they included 362 intraresidue, 260 sequential, and 34 long-range restraints. This set of distance restraints has an average flat-well width (difference between upper and lower bounds) of 1.5 Å, with a minimum of 0.11 Å for geminal distances H2′−H2″ and maximum of 4.19 Å for the H3′T11−M7T12 distance (distance bounds 3.27−7.46 Å). The flat-well width of a particular distance depends on the NOE intensity of the corresponding cross-peak (compared to the random noise level and integration error), on NOE intensities involving all surrounding protons, and on self-consistency of the data. For example, for our data, distances with the center of the flat well between 3.0 and 4.0 Å had an average flat-well width of 1.75 Å.

The distances derived from 2D NOESY spectra acquired in D_2O were further supplemented with 204 distance restraints involving exchangeable protons based on the data from the H_2O NOESY spectrum at mixing time 300 ms (categorized as weak, medium, and strong) and 108 Watson−Crick hydrogen bond restraints (1.7−1.9 Å for distances involving protons and 2.8−2.9 Å for distances involving heavy atoms). In total, this constituted 968 distance restraints, or 19.4 restraints/residue. This set of restraints was used for the quality evaluation of the refined structures (calculation of residual distance violation and NOE-based *R*-factors). However, a large percentage of these distances, such as fixed distances or intrasugar distances with little conformational dependence, does not have much structural information; they were excluded from the set used in the refinement. The final set of restraints used for the refinement consisted of 628 lower and upper distance bounds, or 12.6 per residue: 376 distances involving nonexchangeable protons (124 intraresidue, 218 sequential, and 34 long-range), 144 distances involving exchangeable protons (sequential or long-range), and 108 distances corresponding to the Watson−

Crick hydrogen bonds for the base-paired residues observed in the H_2O NOESY spectrum.

The refinement protocol involved a series of DYANA, DNAminiCarlo, and AMBER calculations. For the DYANA calculations, weak torsion angle restraints for the helical segments and very strong (weight of 100.0) bond length restraints for the C4′−O4′ bonds in sugars were included as well. Two hundred random structures were generated and subjected to simulated annealing in torsion angle space with DYANA. Some of the resulting structures had incorrect prochiral centers; they were discarded from further consideration.

Structures with the best score and correct prochiral centers were further refined with DNAminiCarlo. For that purpose, helical parameters were calculated for the structures produced by DYANA and used as input for DNAminiCarlo. As the first step, the refinement was carried out separately for the two domains of PSQ: one consisting of a dimer of residues G1 through C4 and G16 through C25 which constituted a continuous double helix (14 base pairs) with a nick between C4 and G16, and another consisting of a dimer of residues C5 through T15 constituting another double helix (six base pairs) with single stranded loops T11−T15. For each of the domains, the dyad symmetry was imposed during the refinement. For the second step, the two domains were assembled together and further refined with DNAminiCarlo. There are no observed NOE cross-peaks between the two domains, and this step of refinement was mainly driven by the force field which affected mostly only their mutual orientation.

Finally, the structures were restrained minimized with AMBER 4.1 in vacuo with the Cornell et al. force field (*31*) using 200 steps of steepest descent and 800 steps of conjugate gradient minimization. The refined structures were deposited in the Brookhaven Protein Data Bank, PDB entry 1SLS, models 1−8. In addition, the structure used for the initial MARDIGRAS calculations was further refined using the final set of distance restraints (PDB 1SLS, model 9). For each of the refined structures, the NOE-based figures of merit were calculated with the CORMA program (*32*).

RESULTS AND DISCUSSION

Divalent magnesium was required for structure stabilization; 1D ^{1}H-NMR spectra show multiple sets of signals in the absence of Mg^{2+}, and the oligonucleotide melts non-cooperatively at these conditions (data not shown). In contrast, the melting is cooperative in the presence of Mg^{2+}. HPLC with the size exclusion column showed that SLS-25 is almost exclusively dimeric at room temperature in the presence of Mg^{2+} (data not shown); there is a single set of ^{1}H resonances at low temperature, indicating that SLS-25 forms a symmetric homodimeric structure with identical conformations of the two strands at these conditions. Interestingly, multiple species were also observed for certain residues (e.g., A7, T8, A20, and C21) at elevated temperature even in the presence of Mg^{2+}. The species are in slow exchange on the NMR time scale; the population of the second species increases with temperature, equaling the population of the first at ∼25 °C (data not shown). The nature of the second conformer is unknown; all subsequent NMR experiments were performed at 10 °C where only the single species was apparent.

Chemical probing with $KMnO_4$ showed that T11−T15 (loops $\mathbf{t_1}$ and $\mathbf{t_2}$) are accessible to modification, while T3, T8, and T17 are protected due to base pairing (Figure 3A). This base pairing was confirmed by the 2D NOE spectrum of SLS-25 in water (Figure 3B). Imino protons are clearly identified in this spectrum for all base pairs expected for dimeric SLS-25, both in the SLS and PSQ conformations. It is interesting that the sequence of the segment **b** in SLS-25 (C5-C6-A7-T8-G9-G10, Figures 1 and 2) is the same as in the immobile 55 residue SLS construct used for the low-resolution studies earlier. The assignment for the T8 imino proton H3, 13.70 ppm (Figure 3B), confirms the assignment for the tertiary AT base pair in the 55 nucleotide construct made previously based on its chemical shift alone, and thus it confirms the SLS formation for that construct (*14*).

All helical segments in SLS-25 display connectivities typical of B-DNA (summarized in Figure 4). Two disconnected helical segments were observed for SLS-25: C5-C6-A7-T8-G9-G10 and G1-C2-T3-C4-G16-T17-G18-C19-A20-C21-G22-A23-G24-C25. Nontrivial features of the stacking interaction are the absence of NOE cross-peaks between residues C4 and C5 and the long-range cross-peaks observed between base pairs C4:G22 and G16:C21 (e.g., G22H1−G16H1, C4H6−G16H8, C4H1′−G16H8, and $C4NH_2$−$C21NH_2$); some of these cross-peaks are shown in Figure 3C. Note that the secondary structures of both SLS and PSQ are consistent with two distinct tertiary structures characterized by different stacking interaction patterns. These two putative conformers are shown schematically in Figure 1A for the SLS; a scheme for the two PSQ conformers is similar (not shown). Our earlier theoretical calculations (*14*, *16*) predicted that only one conformer is possible for SLS, the one corresponding to the left direction of "strand slippage" (Figure 1A). This prediction is confirmed by the observed NOE connectivities for SLS-25. Indeed, for the alternative conformer with the right direction of the "strand slippage" (Figure 1A), one expects connectivities between residues C4 and C5 and the absence of connectivities between C21 and G22 and between C4 and G16.

Nevertheless, the observed stacking interactions are still consistent with both folding isomers for SLS-25, linear SLS, and PSQ (Figure 2, top left). The observed connectivity C4H6−G16H8 (dashed lines in Figure 2, right) can be crucial for distinguishing between the two isomers. In linear SLS, this interaction is intramolecular. In PSQ, this interaction is intermolecular, i.e., between C4H6 and G16′H8, symmetrically related to G16H8. It would be possible to distinguish between the intra- and intermolecular NOE cross-peaks, if we could exchange both protons, C4H6 and G16H8, for deuterium. Because only H8 protons of purines are easily exchanged (*33*), we synthesized a related oligonucleotide, SLS1−25, with the C4:G22′ base pair substituted by G4:C22′ (Figure 2, lower left). SLS1−25 and SLS-25 have similar proton chemical shifts (one residue apart from the substituted nucleotides) and the same connectivities, including the G4H8−G16H8 NOE (data not shown). This indicates that SLS-25 and SLS1−25 structures must be similar as well.

The SLS1−25 sample, which had H8 positions of purines exchanged for deuterium, was annealed in a 1:1 ratio with a fully protonated SLS1−25 sample, so that a combinatorial mixture of fully protonated, fully deuterated, and protonated/

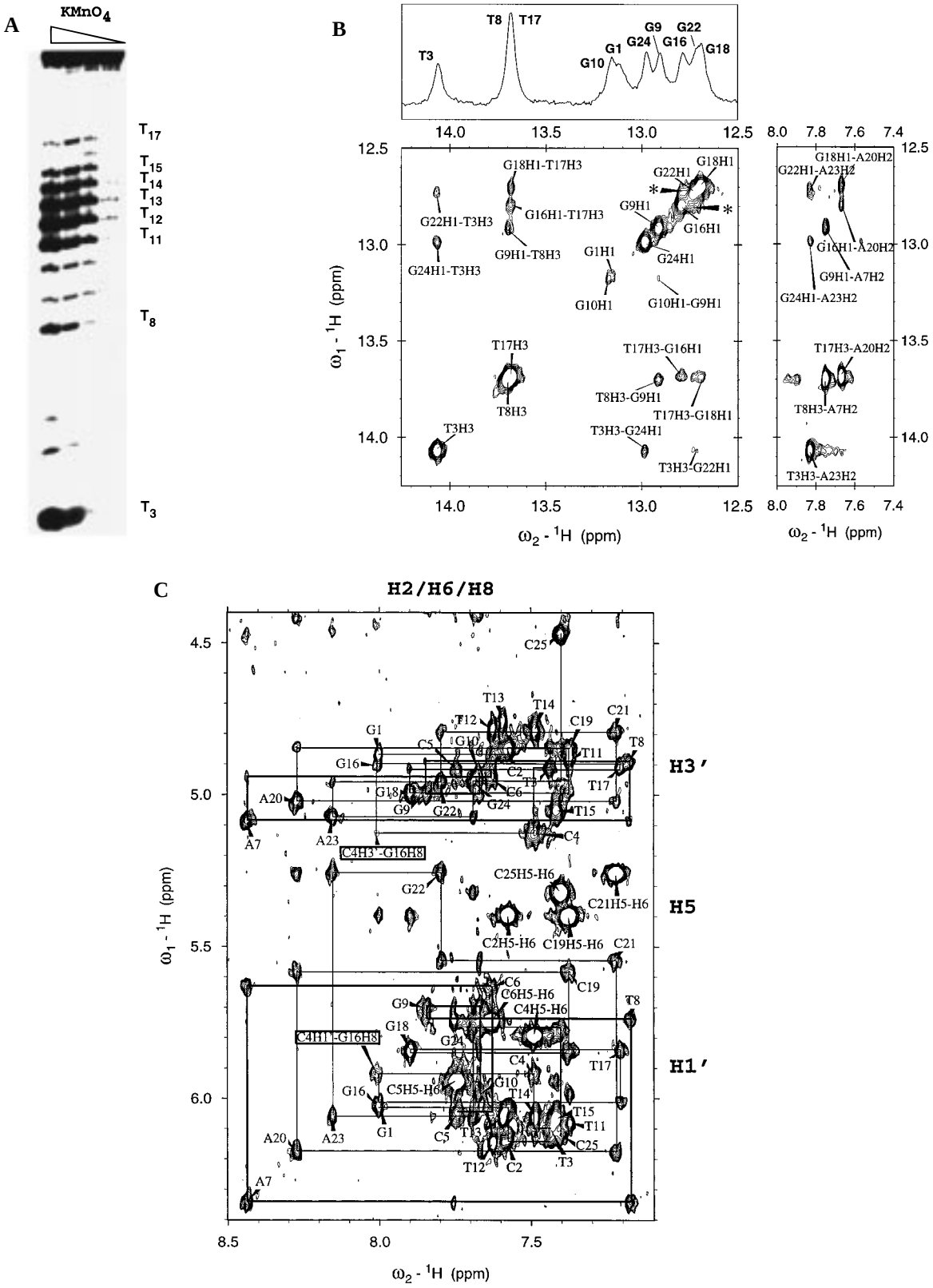

FIGURE 3: (A) Polyacrylamide gel electrophoresis of ^{32}P 5′-end-labeled SLS-25 modified by $KMnO_4$ and cleaved in hot piperidine. Concentration of $KMnO_4$ in the four lanes (from left to right) was 0.4, 0.2, 0.04, and 0.02 mg/mL, respectively. Note that residues T11−T15 (loop $\mathbf{t_1}/\mathbf{t_2}$) are accessible, while T3, T8, and T17 are protected. (B) Two lower panels show portions of SLS-25 600 MHz NOESY spectrum in water at 10 °C (mixing time 300 ms) with imino proton assignments; the upper panel shows 1D SLS-25 spectrum in water with 1331 water suppression. Asterisks denote long-range NOE cross-peaks G16H1−G22H1 which are not well resolved due to proximity to the diagonal. These cross-peaks, as well as a number of others between base pairs C4:G22 and G16:C21, prove the type of stacking interaction corresponding to the left direction of strand slippage shown in the left panel of Figure 1A. (C) A portion of SLS-25 600 MHz NOESY spectrum in D_2O at 10 °C, mixing time 75 ms. Base proton/H1′ and base proton/H3′ connectivities are shown for helical segments; intraresidue cross-peaks are labeled. Connectivities for the G1-C2-T3-C4-G16-T17-G18-C19-A20-C21-G22-A23-G24-C25 segment are shown in thin lines, and those for the C5-C6-A7-T8-G9-G10 are shown in thick lines. Tertiary cross-peaks C4H1′−G16H8 and C4H3′−G16H8 are shown with boxed labels.

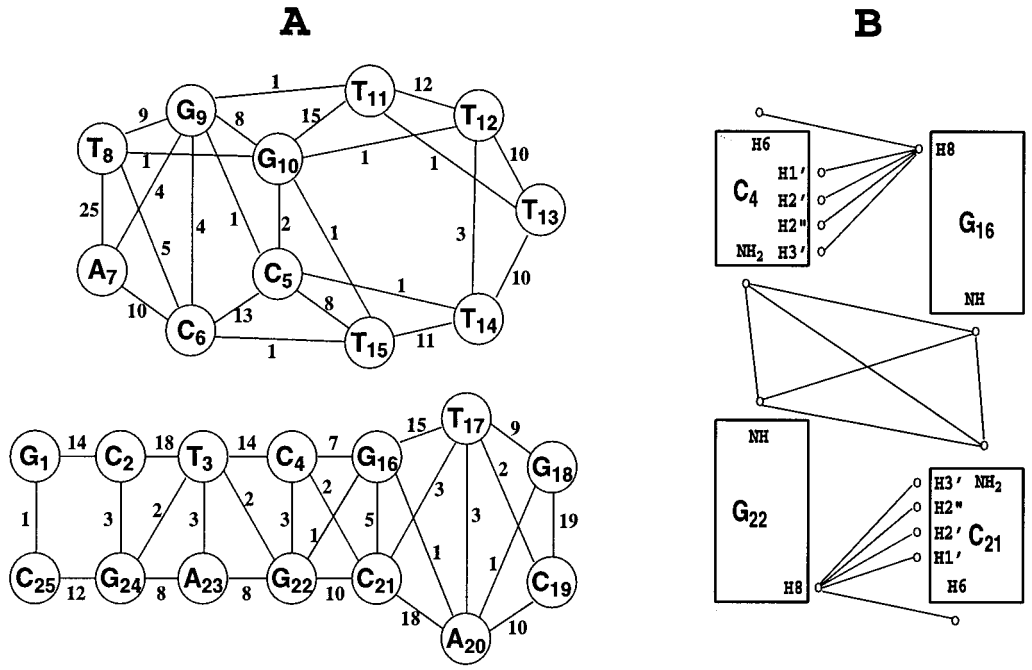

FIGURE 4: (A) Summary of interresidue NOE connectivities for SLS-25. Lines are drawn between the residues connected by NOE cross-peaks; numbers of cross-peaks are shown. Note that there are two domains in SLS-25 with no observed NOE interactions between them. The total number of observed cross-peaks is greater than the number of distance restraints used for the structure determination, because not all NOE volumes could be reliably integrated, and certain distance restraints were not used for the refinement (see Materials and Methods). (B) Summary of the most important observed tertiary interactions between base pairs C4:G22 and G16:C21. This figure was prepared with the help of the Sparky program (*23*).

deuterated SLS1−25 dimers ensued in the ratio 1:1:2. Intensities of NOE cross-peaks in the NOESY spectra of the fully protonated SLS1−25 and of the protonated/deuterated mixture were normalized using cross-peaks not involving exchangeable H8 protons as a reference. After normalization, intensities of all cross-peaks involving a single H8 proton in the protonated/deuterated mixture were 62 ± 14% of the corresponding peaks from the fully protonated sample (averaged over 51 H8 cross-peaks). The expected value is 50% (because half of the H8 protons in the mixture must be exchanged for deuterium); the discrepancy might be due to incomplete deuterium exchange at some H8 positions or due to a small excess of protonated sample in the mixture. The intensity of the G10H8−G9H8 cross-peak (Figure 5) also lies in this range: its intensity in the mixed sample constituted 57% of that in the fully protonated SLS1−25. This is expected of intrastrand cross-peaks, because only half of the strands are deuterated. However, the NOE intensity of the G4H8−G16H8 cross-peak decreased to 27% in the mixed sample compared to the fully protonated SLS1−25 sample (Figure 5). This is possible only if the G4H8−G16H8 cross-peak is due to intermolecular interaction, i.e., G4H8−G16′H8, meaning that the kissing isomer (PSQ) is adopted under the conditions used. Indeed, ^{1}H-NMR signals come only from species with both positions protonated; for an intermolecular NOE, both strands would have to be protonated, thus reducing the expected value to 25%. The discrepancy (observed 27% vs expected 25%) may be due to the reasons discussed above or to integration error. If this NOE was intramolecular, the signal would come also from species with only one strand protonated and would have reduced the NOE to 62 ± 14% (the range estimated from cross-peaks involving a single H8 proton). While we cannot directly estimate the integration error of the G4H8−G16H8 cross-peak, it would need to be >100% to invalidate the conclusion about the interstrand nature of this interaction; we consider such an error unlikely. Furthermore, analysis of the refined PSQ structure provides additional arguments in favor of this conclusion (vide infra).

Refinement of the PSQ structure for SLS-25 was carried out as described in the Materials and Methods section. The refined structures are shown in Figure 6, and structural statistics are given in Table 1. The structures were refined to a relatively low value of R^x-factor (ca. 7%) and residual distance deviation (0.13 Å). Note that the residual distance deviation was calculated using a full set of 968 distance restraints, while the refinement was carried out with the reduced set of 628 restraints (see Materials and Methods). Most of the omitted restraints provided little structural information; however, some restraints were omitted because we initially considered them unreliable. For example, cross-peaks such as C6H1′−T8M7, G1H1′−T3M7, and G18H1′−C19H1′ were clearly observed mostly due to spin diffusion, which is strong because of the size of a molecule and the high viscosity of D_2O at 10 °C. The MARDIGRAS-derived bounds for these three distances were 4.30−6.56, 3.71−6.66, and 2.89−5.40 Å, respectively. They were not used during the refinement, because we were not convinced that these bounds were accurate enough. However, the resulting distances in the refined structure (6.71, 6.63, and 5.35 Å, respectively) are nearly within the calculated bounds. This argues that random-error MARDIGRAS is capable of calculating accurate bounds even when the NOE is determined largely by spin diffusion. As previously noted, this is possible since the MARDIGRAS algorithm establishes an interproton distance on the basis of all cross-peak intensities and not just the cross-peak associated with that pair of protons (*28*).

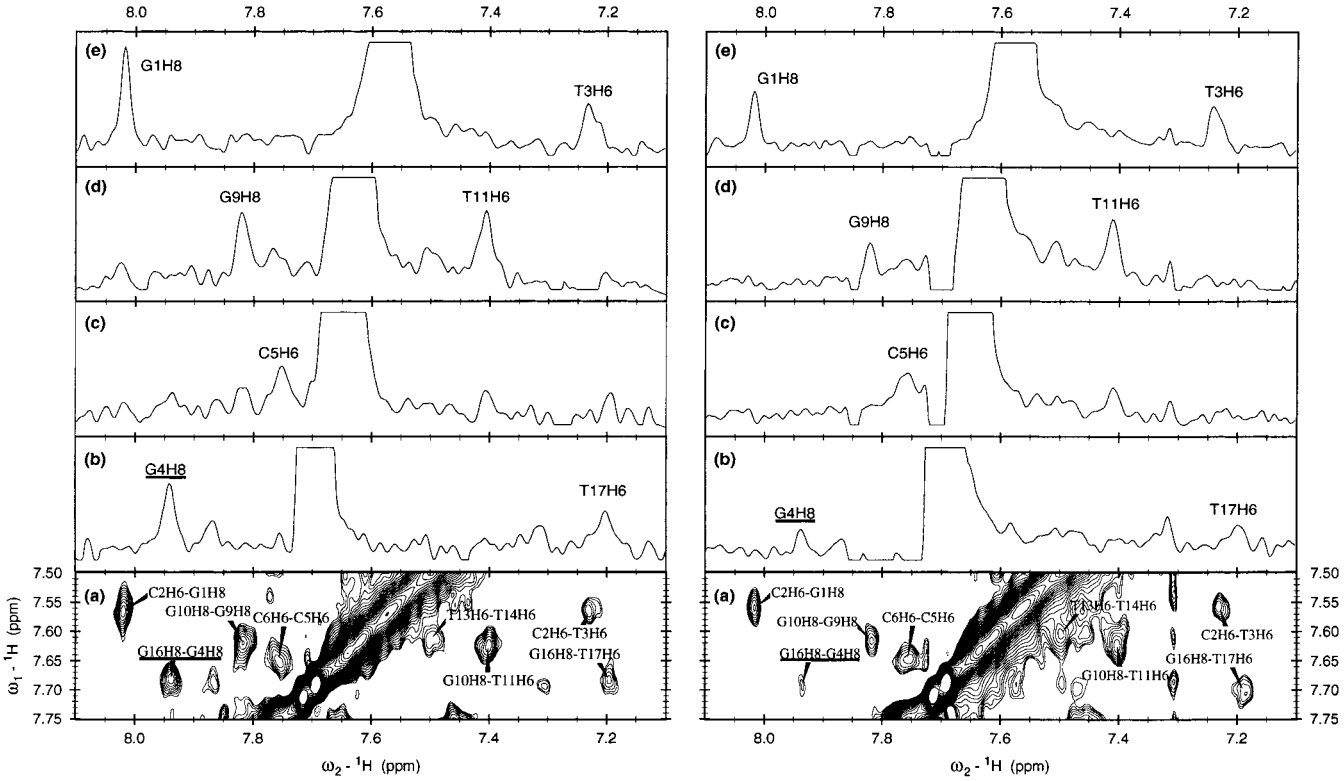

FIGURE 5: Portions of 500 MHz 2D NOESY spectra of SLS1−25 in D_2O at 10 °C, mixing time 300 ms (a), and 1D cross-sections of these spectra through $\omega 1$ = G16H8 (b), C6H6 (c), G10H8 (d), C2H6 (e). Left panels show spectra of fully protonated SLS1−25; right panels show spectra of protonated/deuterated mixture of SLS1−25 (see text). The 1D slices were normalized in such a way that cross-peaks between two nonexchangeable protons (e.g., C2H6−T3H6, C6H6−C5H6) have the same amplitude in both spectra. Note that cross-peaks involving one exchangeable and one nonexchangeable proton (e.g., C2H6−G1H8 and G16H8−T17H6) and cross-peaks between two exchangeable protons from the same strand (G10H8−G9H8) have amplitude decreased by ca. 50% in the spectra of the protonated/deuterated mixture (right panel). At the same time the amplitude of the G16H8−G4H8 (label underlined) is decreased in the $^1H/^2H$ mixture to ca. 25% of that in the fully protonated sample, which proves that this cross-peak is intermolecular, i.e., between G4H8 and G16′H8 (symmetrically related to G16H8).

Table 1: Structural Statistics[a]

number of distance restraints	
nonexchangeable	656
exchangeable	204
holonomic hydrogen bonds restraints	108
total	968
per residue	19.4
energy (kcal/mol)	−1999 ± 39
residual distance deviation (Å)	0.13 ± 0.01
NOE-based R^x factor ($\times 10^2$)	
75 ms NOESY dataset	6.98 ± 0.19
150 ms NOESY dataset	7.12 ± 0.19
atomic root-mean-square deviation (Å)	
all residues	3.79 ± 1.84
residues 1−4 and 16−25	1.31 ± 0.44
residues 5−10	0.73 ± 0.39

[a] Average parameters for the nine refined SLS-25 models are given, ± standard deviations. Energy was calculated using the AMBER 4.1 force field (*31*) with the AMBER program (*35*); R^x is the sixth-root weighted NOE-based *R*-factor (*39*) calculated with CORMA (*32*).

Table 2: Interproton Distances[a]

assignment[b]	bounds[c]	PSQ distance[d]	SLS distance[e]
T15H3′−C5H5	2.95−5.06	18.84	10.81
T15H6−C5H5	2.64−5.13	19.48	9.79
T15H6−C5H6	2.62−4.83	18.83	7.90
T15M7-C5H5	3.59−6.54	18.60	9.64
T15M7-G10′H1	1.80−6.00	18.96	8.42
T15M7-C5H41	1.80−6.00	19.04	10.27
T15M7-C5H42	1.80−6.00	18.98	10.60
T14M7-C5H42	1.80−6.00	21.17	14.06
T15H3′−C5′H5	2.95−5.06	4.87	16.34
T15H6−C5′H5	2.64−5.13	4.39	12.94
T15H6−C5′H6	2.62−4.83	4.97	13.45
T15M7-C5′H5	3.59−6.54	3.48	11.43
T15M7-G10H1	1.80−6.00	5.12	12.32
T15M7-C5′H41	1.80−6.00	4.05	12.75
T15M7-C5′H42	1.80−6.00	3.50	11.83
T14M7-C5′H42	1.80−6.00	6.14	7.79

[a] Interproton distances corresponding to observed NOE cross-peaks placing the loop residue T15 close to the stem residue C5. [b] The top half of the table lists assignments placing T15 close to C5, and the bottom half lists assignments T15 close to C5′ (symmetrically related to C5); see Figure 7. [c] Distance bounds for nonexchangeable protons were calculated with MARDIGRAS; distance bounds involving one exchangeable proton were based on categorizing the corresponding cross-peaks as "weak". [d] Distances in the NMR structure of the PSQ isomer (PDB 1SLS, model 1). The structure is consistent with the bottom set of the assignments. [e] Distances in the model for the SLS isomer calculated with the DNAminiCarlo program as described before (*14*). Neither set of assignments is consistent with this isomer.

Note that each helical segment is defined to relatively high precision but not the angle between the two (Figure 6, panels A and B). Indeed, residues C5−T15 do not have cross-peaks with the rest of the molecule; therefore, the angle between the two domains is defined mainly by the force field, including the backbone closure requirements and, indirectly, by the conformation of the loop T11−T15. All structures with best scoring functions have a roughly parallel orientation of the two domains (PDB entry 1SLS, models 1−8, and Figure 6C), except one (model 9) which has an almost perpendicular orientation of the domains (Figure 6d). While

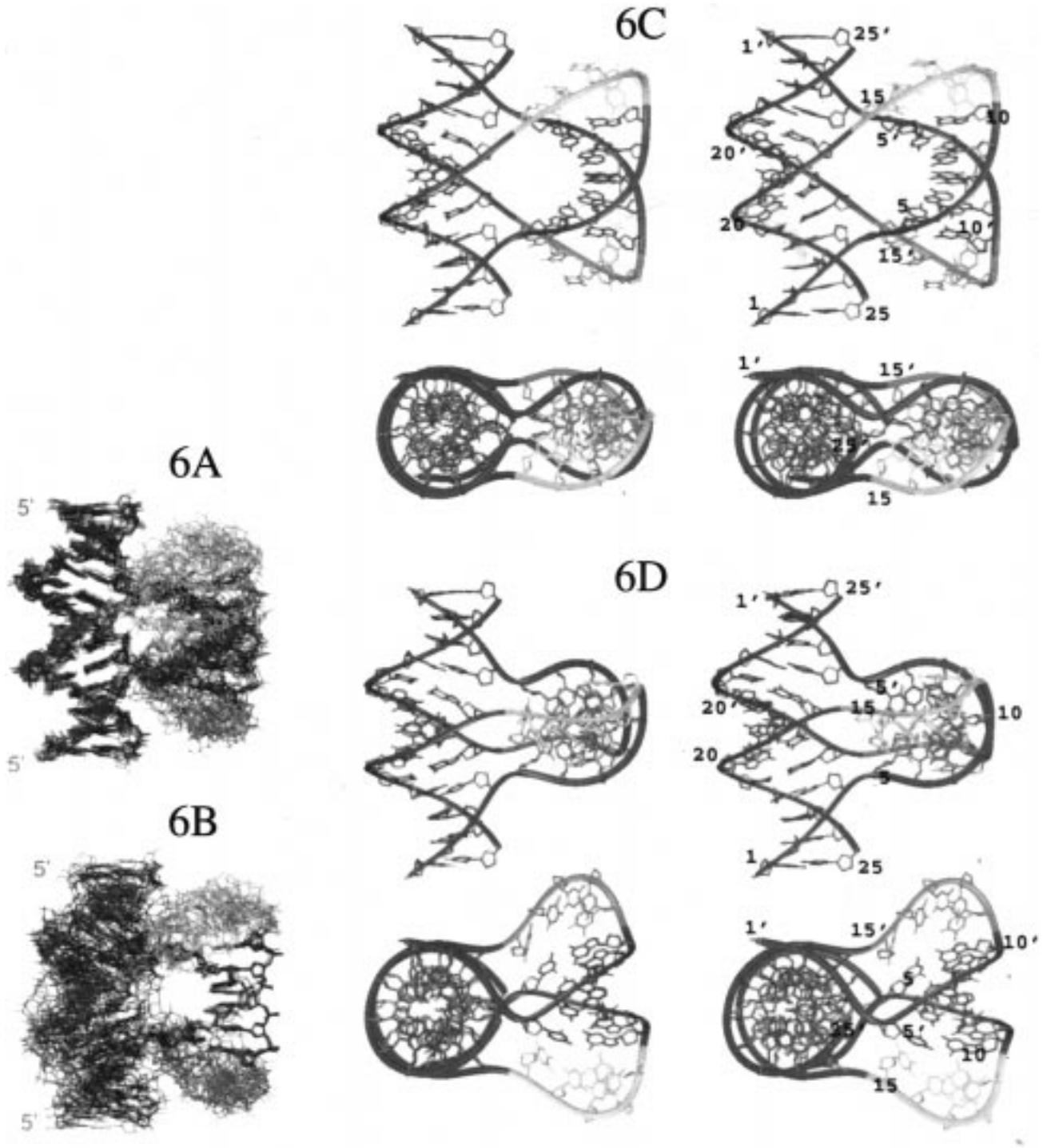

FIGURE 6: NMR structures of SLS-25 pseudosquare knot; the coloring scheme is the same as in Figure 2, right. MidasPlus program (*40*) was used to generate this figure and Figure 7. (A) Superpositions of nine refined SLS-25 structures. Residues 1−4, 16−25, 1′−4′, and 16′−25′ were used for superposition. (B) Residues 5−10 and 5′−10′ were used for superposition. (C) Stereopair of the best SLS-25 structure (model 1) with parallel orientation of the two domains; two perpendicular views. (D) Stereopair of SLS-25 model 9 with perpendicular orientation of the two domains; two perpendicular views. Note the greater displacement of base pairs from the helical axis (top view) compared to the model 1; this is responsible for the perpendicular orientation of the domains.

this structure has, on average, 70 kcal/mol higher conformational energy, the residual distance deviations and NOE-based R^x-factors are almost the same for all nine models. Because of the short range of NMR-derived distances, the angle between the two domains remains undetermined. In general, this angle appears to be sensitive to the exact geometry of helical grooves (compare Figure 6, panels C and D); a perpendicular orientation of the domains is expected for RNA SLS with A-conformations of the helical segments (*16*). We plan to use the fluorescence resonance energy-transfer technique to resolve this problem, as in the case of an immobile Holliday junction (*34*).

Conformations of the single-stranded loops are also quite variable in the nine refined structures (Figure 6), even though there are enough observed NOE cross-peaks for this region (Figure 4). These loops are probably flexible: most residual distance restraint violations come from these regions, which is indicative of conformational averaging (*35*). It is interest-

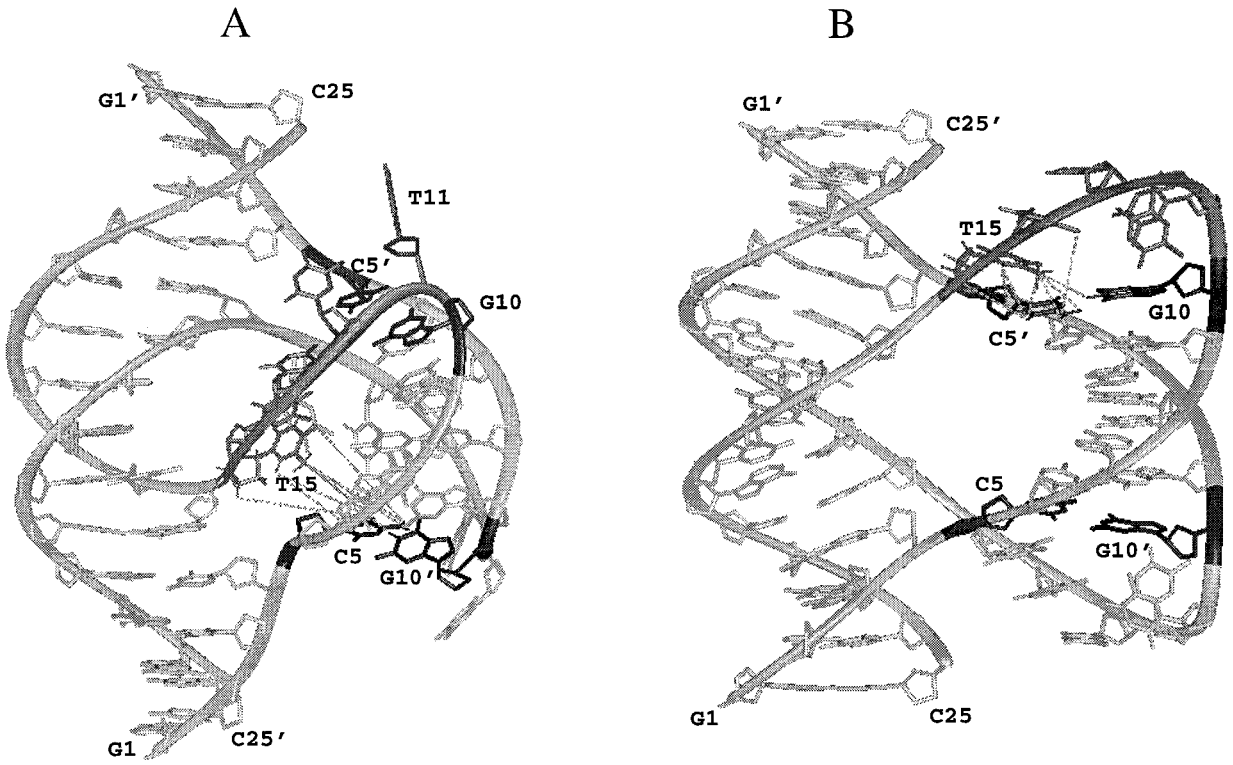

FIGURE 7: Distances placing residue T15 close to residue C5. Loop residues T11 through T15 are shown in dark gray, and residues C5, C5′, G10, and G10′ are shown in black. (A) Model of the SLS isomer calculated as described before (*14*). (B) NMR structure of the SLS-25 PSQ isomer.

Table 3: Helical Parameters for the Stem Regions of the PSQ NMR Structure[a]

base pair step	twist	tilt	roll	shift	slide	rise
G1-C2:G24-C25	36.6 (1.4)	−0.5 (0.9)	6.5 (1.7)	−0.40 (0.12)	−0.61 (0.12)	3.45 (0.09)
C2-T3:A23-G24	30.9 (1.0)	1.4 (0.9)	9.9 (2.9)	0.43 (0.17)	−0.17 (0.16)	3.31 (0.05)
T3-C4:G22-A23	38.3 (2.2)	−0.3 (1.6)	14.5 (3.6)	−0.48 (0.18)	0.35 (0.28)	3.41 (0.18)
C4-G16:C21-G22	39.0 (2.2)	−5.9 (2.1)	5.5 (3.0)	−0.59 (0.34)	0.86 (0.51)	2.96 (0.11)
G16-T17:A20-C21	36.1 (2.4)	5.2 (1.5)	12.7 (2.8)	0.52 (0.22)	−0.09 (0.05)	3.36 (0.12)
T17-G18:C19-A20	29.0 (1.1)	−0.5 (0.9)	16.4 (2.0)	0.29 (0.12)	−0.25 (0.25)	3.22 (0.07)
G18-C19:G18-C19	36.4 (0.8)	0.2 (0.3)	16.7 (4.2)	0.03 (0.03)	−0.28 (0.12)	3.46 (0.11)
C5-C6:G9-G10	37.4 (1.9)	−1.9 (1.4)	−2.0 (2.2)	−0.22 (0.15)	0.04 (0.29)	3.22 (0.13)
C6-A7:T8-G9	36.9 (1.1)	−2.4 (0.8)	8.1 (1.5)	−0.62 (0.10)	0.10 (0.16)	3.31 (0.04)
A7-T8:A7-T8	33.0 (1.2)	0.0 (0.1)	2.8 (1.3)	0.02 (0.01)	−0.65 (0.06)	3.26 (0.05)

[a] Base pair step parameters twist, tilt, roll (degrees), and shift, slide, and rise (Å) are defined according to the Cambridge Convention (*41*) and calculated with the homemade program FITPARAM. Average and standard deviation (in parentheses) values were calculated for the nine NMR structures. Base pair steps for only one-half of symmetric PSQ structure are listed; parameters of individual symmetric steps (e.g., G1-C2:G24-C25 and G1′-C2′:G24′-C25′) were included when calculating average and standard deviation values. Steps which deviated most from the perfect symmetry were T3-C4:G22-A23 and C4-G16:C21-G22 (Δtwist = 4°, Δrise = 0.2 Å).

ing that a number of cross-peaks were observed to place the T15 residue of the loop close to the C5 residue of the secondary stem (Table 2). Based on structure, these cross-peaks must be clearly assigned as T15−C5′, T15−G10, and T14−C5′ for the PSQ isomer (Figure 5B and Table 2). If this were an SLS isomer, they should probably be assigned as T15−C5, T15−G10′, and T14−C5 (Figure 5A); however, these assignments strongly violate experimental bounds (Table 2). When we tried to enforce these restraints by increasing the force constant during refinement of the SLS isomer, it led to severe deterioration of conformational energy without significantly improving distance violations (data not shown). This confirms the assignment of the PSQ isomer based on the deuteration experiment.

The refined PSQ structure has an approximate dyad symmetry and exhibits several interesting features. The total linking number between the two strands is zero, and there is a large cavity between domains. This might explain a relatively low melting temperature (ca. 40 °C, even in the presence of $MgCl_2$). The backbone exhibits sharp turns when connecting the domains, and the interphosphorus distances C5P−G16′P and C5′P−G16P are very short (5.1 and 4.4 Å for structures with parallel and perpendicular orientation of the domains, respectively). These sites are likely candidates for the Mg^{2+}-binding sites. Our preliminary results of the manganese titration of SLS-25 are consistent with this hypothesis; however, additional studies are required. ^{31}P-NMR spectra of SLS-25 were not conclusive at 10 °C because of very broad line width; however, the spread of ^{31}P resonances at elevated temperatures is consistent with the conformation having an irregular backbone (data not shown).

All helical segments have a B-type conformation with some A-like features. Note, for example, that there is a high

positive roll in most base pair steps (Table 3) and that base pairs are somewhat displaced from the helical axis (Figure 6, panels C and D, top views). However, all stem residues have C1′exo/C2′endo sugar puckers. Such a combination of overall B-type geometry with some A-like features is typical of short DNA duplexes in solution (*36*), arguing that the helical segments of PSQ are probably not strained conformationally. In this respect, it is also interesting that the DNA hexamer $(CCATGG)_2$ with the sequence identical to the miniduplex **b:b′** within PSQ SLS-25 has been studied free in solution by NMR previously (*37*). The proton chemical shifts are very similar for the free hexamer and CCATGG within SLS-25, except for terminal residues.

A hypothetical interconversion between the two folding isomers can be achieved without breaking the core of the structure; it only requires a local melting of two terminal stems, **a:d** and **a′**:**d′** and reannealing them into **a:a′** and **d:d′** (Figure 1D). This mechanism might explain the species observed at elevated temperature (see above). Alternatively, that species might be due to a rotation of the PSQ domains relative to each other. In this respect, of note is the irregular shape of NOE cross-peaks involving C4 protons (Figure 3C), which might be indicative of some conformational averaging even at low temperature.

CONCLUSIONS

We described a novel SLS-related folding motif, pseudosquare knot, and determined its solution structure. Even though the data reported here were obtained for DNA oligonucleotides, model-building (*16*) and our preliminary chemical modification data show that SLS can be formed for RNA molecules as well. The question of SLS formation in double-stranded DNA with short direct repeats remains open (see, e.g., ref *2*); this problem has not been addressed in our studies. The strand slippage mechanism in DNA was first proposed based on the S1 nuclease cleavage data for the alternating GA sequences in the sea urchin histone genes (*6*). Even though these original data are best explained by another unusual DNA structure, the H-form triplex (*4*), S1 cleavage data for some other DNA sequences are compatible with SLS but not H-DNA (*7*). However, we see the main significance of our findings in the characterization of the novel tertiary fold for nucleic acids with potential importance for single-stranded genomes, both DNA and RNA. Sequence motifs compatible with formation of immobile SLS PSQ are frequent in single-stranded genomes; such conformations may play a structural and/or functional role. In particular, the pseudosquare knot may form during dimerization of single-stranded nucleic acids, not unlike the "kissing loop" formation during dimerization of retroviral RNA (*38*). Incidentally, we have found an abundance of sequences with a potential to form a homodimeric PSQ or SLS in retroviral genomes (unpublished data). The preference of the PSQ over the linear SLS isomer may be dependent on both sequence and conditions; putative interconversion between isomers may serve as a regulatory switch.

ACKNOWLEDGMENT

We thank V. B. Zhurkin, U. Schmitz, C. Kojima, H. Liu, and P. Lukavsky for many useful discussions. N.B.U. is grateful to U. Schmitz, D. Z. Avizonis, V. J. Basus, A. Mujeeb, L. J. Yao, and M. Gochin for their help and encouragement in this, his first experimental project. We gratefully acknowledge use of the UCSF Computer Graphics Laboratory.

REFERENCES

1. Cantor, C. R., Bondopadhyay, S., Brahmachari, S. K., Hui, C.-F., McClelland, M., Morse, R., and Smith, C. L. (1988) in *Unusual DNA Structures* (Wells, R. D., and Harvey, S. C., Eds.) pp 73−89, Springer-Verlag, New York.
2. Yagil, G. (1991) *Crit. Rev. Biochem. Mol. Biol. 26*, 475−559.
3. Rich, A., Nordheim, A., and Wang, A. H. (1984) *Annu. Rev. Biochem. 53*, 791−846.
4. Mirkin, S. M., Lyamichev, V. I., Drushlyak, K. N., Dobrynin, V. N., Filippov, S. A., and Frank-Kamenetskii, M. D. (1987) *Nature 330*, 495−497.
5. Parsons, C. A., Murchie, A. I., Lilley, D. M., and West, S. C. (1989) *EMBO J. 8*, 239−246.
6. Hentschel, C. C. (1982) *Nature 295*, 714−716.
7. Mace, H. A., Pelham, H. R., and Travers, A. A. (1983) *Nature 304*, 555−557.
8. Gorgoshidze, M. Z., Minyat, E. E., Gorin, A. A., Demchuk, E. Ia., Farutin, V. A., and Ivanov, V. I. (1992) *Molek. Biol. (Eng. Transl.) 26*, 1263−1273.
9. Puglisi, J. D., Wyatt, J. R., and Tinoco, I., Jr. (1990) *J. Mol. Biol. 214*, 437−453.
10. Kim, E. L., Esparza, F. M., and Stachowiak, M. K. (1996) *J. Neurochem. 67*, 26−36.
11. Zhang, S., and Seeman, N. C. (1994) *J. Mol. Biol. 238*, 658−668.
12. Carlström, G., Chen, S. M., Miick, S., and Chazin, W. J. (1995) *Methods Enzymol. 261*, 163−182.
13. Overmars, F. J., Lanzotti, V., Galeone, A., Pepe, A., Mayol, L., Pikkemaat, J. A., and Altona, C. (1997) *Eur. J. Biochem. 249*, 576−583.
14. Ulyanov, N. B., Bishop, K. D., Ivanov, V. I., and James, T. L. (1994) *Nucleic Acids Res. 22*, 4242−4249.
15. Minyat, E. E., Khomyakova, E. B., Petrova, M. V., Zdobnov, E. M., and Ivanov, V. I. (1995) *J. Biomolec. Struct. Dyn. 13*, 523−527.
16. Farutin, V. A., Gorin, A. A., Zdobnov, E. M., and Ivanov, V. I. (1997) *J. Biomolec. Struct. Dyn. 15*, 45−52.
17. Chang, K. Y., and Tinoco, I., Jr. (1997) *J. Mol. Biol. 269*, 52−66.
18. Ulyanov, N. B., Ivanov, V. I., Minyat, E. E., Khomyakova, E. B., Petrova, M. V., Lesiak, K., and James, T. L. (1998) in *Proceedings Tenth Conversation in the Discipline Biomolecular Stereodynamics*, Adenine Press (in press).
19. Jeener, J., Meier, B. H., Bachmann, P., and Ernst, R. R. (1979) *J. Chem. Phys. 71*, 4549−4553.
20. States, D. J., Haberkorn, R. A., and Ruben, D. J. (1982) *J. Magn. Reson. 48*, 286−292.
21. Smallcombe, S. H. (1993) *J. Am. Chem. Soc. 115*, 4776−4785.
22. Day, M. (1993) *Striker*, University of California, San Francisco, CA.
23. Kneller, D., Kuntz, I. D., and Goddard, T. (1996) *Sparky*, University of California, San Francisco, CA.
24. Güntert, P., Mumenthaler, C., and Wüthrich, K. (1997) *J. Mol. Biol. 273*, 283−298.
25. Pearlman, D. A., Case, D. A., Caldwell, J. W., Ross, W. S., Cheatham, T. E., III, Ferguson, D. M., Seibel, G. L., Singh, U. C., Weiner, P., and Kollman, P. A. (1995) *AMBER*, ver. 4.1, University of California, San Francisco, CA.
26. Ulyanov, N. B., Gorin, A. A., and Zhurkin, V. B. (1989) in *Proceedings Supercomputing′89: Supercomputer Applications* (Kartashev, L. P., and Kartashev, S. I., Eds.) pp 368−370, International Supercomputing Institute, Inc., St. Petersburg, FL.
27. Zhurkin, V. B., Lysov, Yu. P., and Ivanov, V. I. (1978) *Biopolymers 17*, 377−412.

28. Borgias, B. A., and James, T. L. (1989) *Methods Enzymol. 176*, 169−183.
29. Liu, H., Spielmann, H. P., Ulyanov, N. B., Wemmer, D. E., and James, T. L. (1995) *J. Biomolec. NMR 6*, 390−402.
30. Ulyanov, N. B., Schmitz, U., and James, T. L. (1993) *J. Biomolec. NMR 3*, 547−568.
31. Cornell, W. D., Cieplak, P., Bayly, C. I., Gould, I. R., Merz, K. M., Jr., Ferguson, D. M., Spellmeyer, D. C., Fox, T., Caldwell, J. W., and Kollman, P. A. (1995) *J. Am. Chem. Soc. 117*, 5179−5197.
32. Keepers, J. W., and James, T. L. (1984) *J. Magn. Reson. 57*, 404−426.
33. Brandes, R., and Ehrenberg, A. (1986) *Nucleic Acids Res. 14*, 9491−9508.
34. Miick, S. M., Fee, R. S., Millar, D. P., and Chazin, W. J. (1997) *Proc. Natl. Acad. Sci. U.S.A. 94*, 9080−9084.
35. Yao, L. J., James, T. L., Kealey, J. T., Santi, D. V., and Schmitz, U. (1997) *J. Biomolec. NMR 9*, 229−244.
36. Ulyanov, N. B., and James, T. L. (1995) *Methods Enzymol. 261*, 90−120.
37. Tran-Dinh, S., Neumann, J.-M., Huynh-Dinh, T., Allard, P., Lallemand, J. Y., and Igolen, J. (1982) *Nucleic Acids Res. 10*, 5319−32.
38. Paillart, J. C., Skripkin, E., Ehresmann, B., Ehresmann, C., and Marquet, R. (1996) *Proc. Natl. Acad. Sci. U.S.A. 93*, 5572−5577.
39. James, T. L. (1991) *Curr. Opin. Struct. Biol. 1*, 1042−1053.
40. Ferrin, T. E., Huang, C. C., Jarvis, L. E., and Langridge, R. (1988) *J. Mol. Graphics 6*, 13−27.
41. Dickerson, R. E., Bansal, M., Calladine, C. R., Diekmann, S., Hunter, W. N., Kennard, O., von Kitzing, E., Lavery, R., Nelson, H. C. M., Olson, W. K., Saenger, W., Shakked, Z., Sklenar, H., Soumpasis, D. M., Tung, C.-S., Wang, A. H.-J., and Zhurkin, V. B. (1989) *J. Mol. Biol. 208*, 787−791.

BI981018D