BIOCHEMISTRY

 Volume 47, Number 18 May 06, 2008

Articles

Bulged Adenosine Influence on the RNA Duplex Conformation in Solution[†,‡]

Lukasz Popenda, Ryszard W. Adamiak, and Zofia Gdaniec*

Institute of Bioorganic Chemistry, Polish Academy of Science, Noskowskiego 12/14, 61-704 Poznań, Poland

Received December 20, 2007; Revised Manuscript Received March 10, 2008

ABSTRACT: The RNA single bulge motif is an unpaired residue within a strand of several complementary base pairs. To gain insight into structural changes induced by the presence of the adenosine bulge on RNA duplex, the solution structures of RNA duplex containing a single adenine bulge (5′-GCAGAA-GAGCG-3′/5′-CGCUCUCUGC-3′) and a reference duplex with all Watson−Crick base pairs (5′-GCAGAGAGCG-3′/5′-CGCUCUCUGC-3′) have been determined by NMR spectroscopy. The reference duplex structure is a regular right-handed helix with all of the attributes of an A-type helix. In the bulged duplex, single adenine bulge stacks into the helix, and the bulge region forms a well-defined structure. Both structures were analyzed by the use of calculated helical parameters. Distortions induced by the accommodation of unpaired residue into the helical structure propagate over the entire structure and are manifested as the reduced base pairs inclination and x-displacement. Intrahelical position of bulged adenine A5 is stabilized by efficient stacking with 5′-neighboring residues G4.

RNA molecules form complex tertiary structures that are closely related with their functions *in vivo*. In addition to standard regular duplexes, there are diverse RNA motifs serving as building blocks of larger RNA tertiary structures. The simplest RNA motifs include mismatches, bulges, base triples, internal and hairpin loops, and dinucleotide platforms. Bulges are formed where double stranded regions of RNA are disrupted by the presence of nucleotide residues that do not have pairing partners in the complementary strand. They are common motifs in folded RNA molecules providing the structural flexibility required for RNA folding. Bulges can also participate in RNA−RNA interactions and RNA−ligand binding (*1*−*5*), and provide sites for specific protein−RNA interactions (*6*−*8*).

A single unpaired base on one strand forms the smallest bulge, and it can either be stacked into the helix or can be directed away from the helix. In the stacked conformation, a bulged nucleotide residue can be positioned between the adjacent base pairs or can form a dinucleotide platform being arranged coplanar to the base, either to the 5′ or 3′ site of the bulge. The knowledge of the preferred conformation of a single-base bulge may be important for the prediction and understanding of the three-dimensional structure of RNAs. The structure of DNA and RNA bulges has been studied by X-ray (*9*−*13*) and NMR[1] methods (*14*−*22*), with the use of algorithms searching the conformational space for energetically favorable structures (*23*) and molecular simulations

[†] This work was supported by a research grant form the Polish Ministry of Science and Higher Education (No. 2 P04 A 033 30).

[‡] Structure coordinates have been deposited in the Protein Data Bank as entries 2JXQ and 2JXS. Chemical shift data have been deposited in BioMagResBank (BMRB), entry codes 15571 and 15572.

* To whom correspondence should be addressed. Phone: (+48-61) 852 85 03. Fax: (+48-61) 852 05 32. E-mail: zgdan@ibch.poznan.pl.

[1] Abbreviations: NMR, nuclear magnetic resonance; NOE, nuclear Overhauser effect; COSY, correlation spectroscopy; NOESY, nuclear Overhauser effect spectroscopy; DQF-COSY, double quantum filtered spectroscopy; TOCSY, total correlation spectroscopy; HSQC, heteronuclear single quantum correlation spectroscopy.

10.1021/bi7024904 CCC: $40.75
Published on Web 04/10/2008

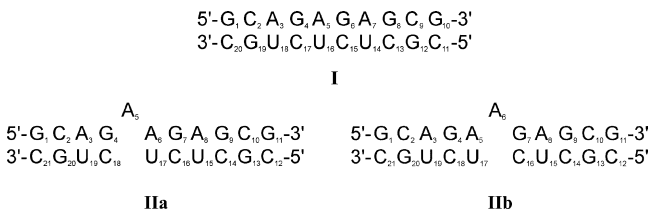

FIGURE 1: The sequence and numbering scheme for the reference (I) and bulged duplexes (II). Two possible secondary structures of the bulged duplex are presented.

(*24*). X-ray crystallography and NMR spectroscopy are the principal methods for the evaluation of structural details of RNAs and their interactions. However, discrepancies appeared for the adenine bulges in the X-ray and NMR determined structures as to whether the adenine bulge is looped out or stacked into the helix. In the high-resolution NMR studies, bulged adenine is generally found stacked into the helix, while the X-ray structures, usually obtained under elevated salt conditions, reveal adenine bulges mostly in looped-out conformation, although intrahelical orientations of the bulge, including dinucleotide platforms, were also identified (*25*).

Despite a number of papers on the nearest-neighbor model for prediction of the free energy of RNA duplexes containing the single nucleotide bulge (*26*−*28*), no relationship has been found yet between the thermodynamics and conformation of the bulge. Distinct sequence features can influence the conformation of the bulge: the type of the bulge, the identity of the neighboring residues, the sequence not adjacent to the bulge, and the length of the helix. In the single nucleotide bulge motif, if two residues of the same type oppose one complementary residue, the bulge can be any of these two identical nucleotides. The RNA sequences with the bulged nucleotide identical to at least one of its neighbors were observed to show enhanced stability, relative to that of the sequences with the nucleotide bulge flanked by residues not identical to the bulge (*26*).

In this work, we have studied the RNA duplex with the 5′-GAAG-3′/5′-CUC-3′ bulge motif. Theoretically, each of the two adenine residues can make a base pair with uridine on the opposite strand, or there might be an equilibrium between these two possible conformations. To get more insight into the structural changes induced by the presence of a single adenine bulge on the duplex, we have determined the structure of the regular duplex, taken as a reference, of identical sequence but missing the bulge. To our knowledge, such a comparative study has not been performed previously for structures with and without bulges. The bulged duplex was designed in such a way that the AA/U motif was inserted between double stranded fragments whose alternate base-pairing scheme was not expected to occur and ensured a good thermodynamic stability. The sequences of both molecules, together with the numbering scheme used in this work, are shown in Figure 1.

MATERIALS AND METHODS

NMR Sample Preparation. Oligoribonucleotides 5′-GCAGAGAAGCG-3′ (A), 5′-GCAGAGAGCG-3′ (B), and 5′-CGCUCUCUGC-3′ (C) were synthesized using standard phosphoramidite chemistry (*29*). Both NMR samples were prepared by mixing equimolar amounts of two strands A + C and B+C. The 1:1 stoichiometric ratio of strands was calculated from molar extinction coefficients. Two RNA strands were dissolved in buffer containing 50 mM NaCl, 10 mM sodium phosphate (pH 6.8), and 0.1 mM EDTA and placed in a Shigemi tube. The duplexes were heated to 80 °C and then allowed to cool slowly to room temperature to achieve annealing. Final RNA concentrations were ∼1 mM. For experiments carried out in D_2O, the duplexes were dried from D_2O three times and redissolved in 99.96% D_2O. A mixture of 90% H_2O and 10% D_2O was used for experiments undertaken to study exchangeable protons.

NMR Spectroscopy. All NMR spectra were collected on a Bruker Avance 600 MHz spectrometer, processed with TopSpin (Bruker Inc.) and analyzed with FELIX (Accelrys) software. Solvent suppression for samples in 90% H_2O /10% D_2O was accomplished using the 3−9−19 WATERGATE pulse sequence (*30*). The residual water peak in D_2O samples was suppressed using low-power presaturation.

Exchangeable proton resonances were observed and analyzed using two-dimensional (2D) NOESY spectra obtained at 10 °C with 150 ms mixing time. The spectra were acquired with a sweep width of 14,000 Hz in both dimensions. Typically, 512 FIDs of 2048 complex points were collected. One-dimensional NOE difference spectra were acquired with selective decoupling of individual resonances during the 2.5 s recycle delay.

NOESY spectra in D_2O were recorded with mixing times of 50, 150, and 400 ms. On average, 2048 complex points in *t2* and 512 FIDs in *t1* were collected within the spectral width of 5000 Hz. For assignment purposes, 2D NOESY spectra in D_2O were collected at temperatures 20, 25, and 30 °C in order to facilitate peak assignments due to slight changes in chemical shift for some resonances. For high resolution DQF-COSY spectra, broadband phosphorus decoupling was achieved by GARP (*31*), and narrow spectral width (2800 Hz) was used for 4096 complex data points. A total of 512 FIDs was collected. TOCSY experiments using the MLEV-17 mixing sequence were performed with a mixing time of 60 ms. Natural abundance ^1H-^{13}C HSQC spectra were recorded within the spectral width of 8500 (^{13}C) and 4000 Hz (^1H) and GARP carbon decoupling during acquisition. Approximately 512 FIDs of 4096 complex points were recorded. Proton-detected ^1H-^{31}P COSY experiments (*32*) were acquired within 2400 Hz spectral width in the ^{31}P dimension and 1500 Hz spectral width in the ^1H dimension; 2048 complex points and 256 FIDs were acquired.

NMR-Derived Restraints. Distance restrains between protons were determined from analysis of the intensities of peaks in the 150 ms NOESY spectrum recorded at 25 °C using the isolated spin approximation. Cross-peak volumes were integrated using the FELIX software package. The average pyrimidine H5−H6 distance (2.45 Å) was used as a reference. The upper and lower bound distance constraints were imposed on the cross-peak intensities to be equal to +30% and −15%, respectively, of the calculated interproton distance.

Watson−Crick base pairs were identified using two criteria: the observation of significantly downfield shifted imino and amino proton resonances and the observation of NOE either for G(N1-H)-C(NH_2) or A(H2)-U(N3-H). The restraints involving Watson−Crick base pairs were based on the standard geometry of nucleic acids and were given a

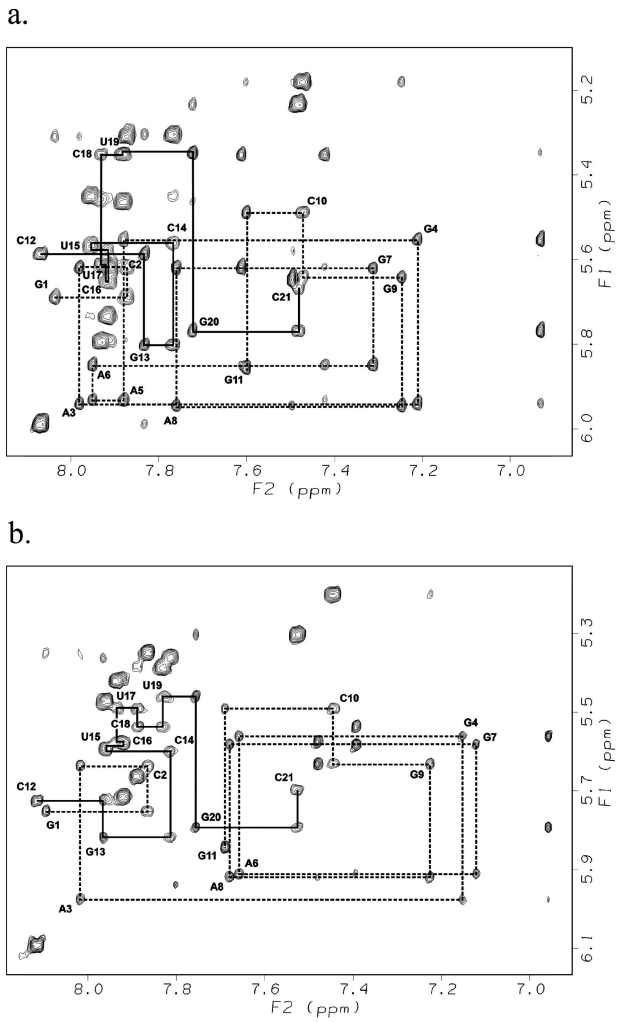

FIGURE 2: Base H6/H8-H1′ regions of the 400 ms NOESY spectra. Sequential connectivities are shown for the top (---) and bottom (—) strands for the reference (a) and bulged duplex (b).

tolerance of ±0.2 Å. For A6-U17 base pair of the bulged duplex, the bounds were given higher tolerance of ±0.4 Å. For all exchangeable protons, the lower limit of NOE distances was set to 1.8 Å and the upper limit to 6.0 Å. Weak base pair planarity restraints (*33*) were applied to the terminal, and all other base pairs for which spectroscopic evidence of pairing was present, except for A6-U17 of the bulged duplex.

Torsion angle restraints were derived from J-coupling measurements of ^{31}P-decoupled ^{1}H - ^{1}H DQF-COSY and ^{1}H -^{31}P COSY cross-peaks (*34*). The torsion angle restraints for sugar rings were determined from an analysis of the 2D DQF-COSY and TOCSY spectra. On the basis of the size of the H1′-H2′ couplings (*34*), which were determined from double quantum filtered COSY experiments, sugar puckers were constrained to be C2′-*endo* or C3′-*endo*. The residues with intermediate H1′-H2′ couplings were loosely constrained to allow for an intermediate sugar puckering.

For both duplexes, the glycosidic torsion angle χ was concluded to be *anti* ($\chi = -158° \pm 30°$) because of the lack of strong intranucleotide H1′-H8 NOE cross-peaks. For C16 and U17 residues of the bulged duplex, the glycosidic torsion angles were left unrestrained because of a strong overlapping of sequential H6-H1′ cross-peaks.

The ^{31}P shifts were used to derive constraints of the backbone torsion angles α and ζ (*35, 36*). When the ^{31}P resonances are in the range typical of helix double-stranded residues, i.e., within the −3.5 to −4.5 ppm, the backbone has the regular helix conformation, and the backbone torsion angles α and ζ have standard helix values (*37*). No downfield shifted phosphorus chemical shifts were found for any residues, and therefore, the angles α and ζ were conservatively restrained to (0° ± 120°) to exclude the *trans* region.

Torsion angle constraints on β (178° ± 30°) were introduced on the basis of the assignment of the ^{31}P signals from 2D spectra and the observation of only very weak connectivities between these ^{31}P resonances and the sequential H5′ and H5″ resonances. The backbone conformation angle ε was constrained on the basis of the $^{3}J_{PH3'}$ couplings estimated from ^{1}H -^{31}P COSY spectra and were given tighter (−153° ± 30°) or looser (−120° ± 120°) constraints for well-resolved and overlapping ^{1}H−^{31}P cross-peaks, respectively (*34*).

The backbone dihedral angle γ was restrained to a range (54 ± 30°) taking into regard the H4′-H5′/H5″ couplings. No backbone torsion angle restraints were imposed on the angles α, β, and γ for A5 and A6 residues in the bulge region. The torsion angle ζ was also left unconstrained for G4 and A5 of the bulged duplex.

Structure Calculation and Refinement. The structures were calculated using the torsion angle molecular dynamics (TAMD) (*38*) algorithm implemented in XPLOR NIH package (*39, 40*). The calculation protocol started from a set of 50 extended structures with proper geometry (bond lengths, bond angles, and impropers) and consisted of four stages: a high-temperature TAMD stage, two cooling steps, one in torsion angle space and one in the Cartesian space, and the final minimization stage. The high-temperature TAMD stage consisted of 8000 steps with a time step of 0.008 ps at a bath temperature of 20,000 K. At this stage, the weight on the distance restraints and van der Waals energy term varied from 2 to 75 kcal mol^{-1} Å^{-2} and from 5 to 0.3, respectively. The energy constant for the experimental dihedrals was set to 100 kcal mol^{-1} rad^{-2}. The high-temperature step was followed by 10,000 cycles of TAMD annealing steps with a time step of 0.007 ps. At this stage the molecules were cooled from 20,000 to 300 K. The NOE force constant was set to 100 kcal mol^{-1} Å^{-2}, and van der Waals forces increased from 0.3 to 1.0. The second cooling stage consisted of 6000 steps in the Cartesian space with temperature decreasing from 3000 to 300 K. The NOE distance force constraint was set to 125 kcal mol^{-1} Å^{-2}, and the dihedral angle constant was set to 200 kcal mol^{-1} rad^{-2} at this stage of the calculations. Afterward, 1200 cycles of the restrained Powell minimization were performed . The calculations at these stages were performed without planarity restrains. Finally, the structures were subjected to the refinement procedure using the torsion angle and base−base positional database potentials of the mean force protocol (*33*). The DELPHIC potential for relative positions of close bases and for torsion angles in nucleic acids was turned on for all residues except A5 of the bulged duplex. Both parameter and topology files used for computations (J. Rife, unpublished results) were derived from the heavy-atoms only topology/parameter set described recently (*41*). The structures with no violation to NOE distances (0.3 Å), dihedral (5°),

Table 1: Structure Determination Statistics[a]

	bulged duplex	reference duplex
distance restraints	337	367
intranucleotide NOEs	196	216
internucleotide NOEs	141	151
negative NOE	3	0
hydrogen bonds	60	60
planarity restraints	9	10
dihedral angle restraints	213	212
backbone	89	92
ribose pucker	105	100
glycosidic	19	20
NOEs per residue	16.05	18.35
NOEs and dihedrals per residue	26.19	28.95
average rms deviation from ideal covalent geometry		
bond length (Å)	0.00140 (0.00001)	0.00134 (0.00003)
angle (°)	0.616 (0.004)	0.610 (0.003)
impropers (°)	0.275 (0.006)	0.255 (0.003)
number of NOE violations (>0.3 Å)	0	0
number of dihedral violations (>5°)	0	0
average non-hydrogen atom pairwise rmsd	0.62 (0.26)	0.30 (0.07)

[a] Standard deviations from the family of structures calculated are given in parentheses.

and with the lowest energies were selected. Helical parameters were calculated with CURVES 5.3 software (*42, 43*).

RESULTS

Assignment of Protons in the Reference Duplex. The NOESY spectra of the reference duplex displayed the characteristic connectivities of a right-handed RNA duplex with all the nucleoside residues in the *anti* conformation. The assignment of nonexchangeable protons was performed using standard methods (*34, 35, 37*). A fragment of 400 ms mixing time NOESY spectrum of the reference duplex is shown in Figure 2a. The aromatic-H1′ assignment pathways are shown by a broken line for the top strand and a solid line for the bottom strand. The absence of all but one H1′-H2′ cross-peak in the DQF-COSY spectra indicated that all residues except the terminal G10 adopt a predominantly N-type sugar pucker.

The NOESY spectrum acquired in 90% H_2O/10% D_2O at 10 °C showing typical Watson−Crick connectivities was employed in the assignment of the exchangeable protons. Uridine imino protons were assigned on the basis of the observation of a strong NOE to the nonexchangeable H2 proton from the adenine residue with which it forms a Watson−Crick base pair. The guanine imino protons were identified on the basis of observable NOEs to the amino protons of hydrogen-bonded cytosine residues. Imino protons of all but terminal residues were assigned and revealed standard intra- and interstrand correlations.

The NMR spectra were consistent with the unperturbed A-RNA structure. No unusual NOEs were observed that might suggest a nonstandard conformation.

Assignment of Protons in the Bulged Duplex. In the bulged duplex, each of two adenine residues, A5 or A6, may make a base pair with U17 on the opposite strand, or an equilibrium between these two possible conformations A5-U17 and A6-U17 should be taken into account (IIa and IIb, Figure 1). The imino (10−15 ppm) and amino (6−9 ppm) regions of ^{1}H NMR spectra recorded in 90% H_2O/10% D_2O provide information on the hydrogen bonding between base pairs. The ^{1}H NMR spectra of bulged duplex show seven sharp and three broad resonances in the region typical of Watson−Crick base pairs. The sharp exchangeable imino resonances were assigned from an analysis of the 2D NOESY spectrum recorded at 10 °C using the methods described above (for the assignments see Supporting Information). Imino protons that do not participate in base pairing resonate between 10.0 and 11.0 ppm. A weak U NH resonance with a chemical shift of 13.94 ppm, corresponding to Watson−Crick base paired uridine, was assigned to the remaining uridine residue, U17. The U17 imino proton is broad and does not yield cross-peaks in the 2D NOESY spectrum. In order to determine which of two possible base pairs, A5-U17 or A6-U17, is formed in the duplex and which of the two neighboring adenine residues remains unpaired, we performed a 1D-NOE difference experiment. After selective irradiation of the imino U17 proton, the NOE effect was observed exclusively to A6-H2 resonance. No NOE was detected to A5-H2, which clearly indicated that only the A6 residue forms a base pair with U17. The broadening of the U17 imino proton might result partially from the dynamical character of the A6-U17 base pair or from higher accessibility of water in the bulge region. In the spectra recorded in D_2O, we did not observe any exchange cross-peaks that could be attributed to a slow equilibrium between A5-U17/A6-U17 base pairs. Nonexchangeable H2/H8 proton resonances of the residues A5, A6, and H5/H6 resonances of U17 in the bulge region are of normal intensities. Additionally, correlations involving these protons in ^{1}H-^{13}C HSQC spectra reveal no broadening as well. If U17 was exchanging between A5 and A6 base pairs, a weak NOE to A5-H2 or a broadening of the resonances involved in the exchange process would be expected.

The NOESY spectra of the bulged duplex in D_2O exhibit typical features of double stranded sequential connectivities. A part of the spectrum is shown in Figure 2b. A continuous set of sequential H8/H6-H1′ NOE connectivities including G4, A5, and A6 residues can be seen in the NOESY spectrum for the upper strand of the bulged duplex. Unfortunately, a continuous determination of the H8/H6-H1′ NOE pathway for the lower stem was not possible because of the cross-peaks overlapping at the C16-U17 step. First, standard proton-based methods were employed to obtain the assignments of the remaining nonexchangeable protons (*34, 35, 37*). These assignments were further confirmed by an analysis of ^{1}H-^{13}C HSQC and ^{1}H-^{31}P HSQC spectra.

As revealed by the absence of most of H1′-H2′ cross-peaks in the DQF-COSY spectra, the residues exhibit mostly N-type sugar pucker. Sugars in the bulge region corresponding to A5 and A6 residues have H1′-H2′ coupling of 3 Hz, which is consistent with a minor contribution of the C2′-*endo* conformation. These residues were loosely constrained during the structure calculation to allow for an intermediate sugar puckering.

The conformation of the bulged residue A5 follows from the NOE contacts observed. A few characteristic cross-peaks have been identified suggesting that the base stacking within the bulge region is uninterrupted and that the bulged residue is accommodated between the adjacent base pairs. First of all, the observation of conventional aromatic H8/H6-H1′

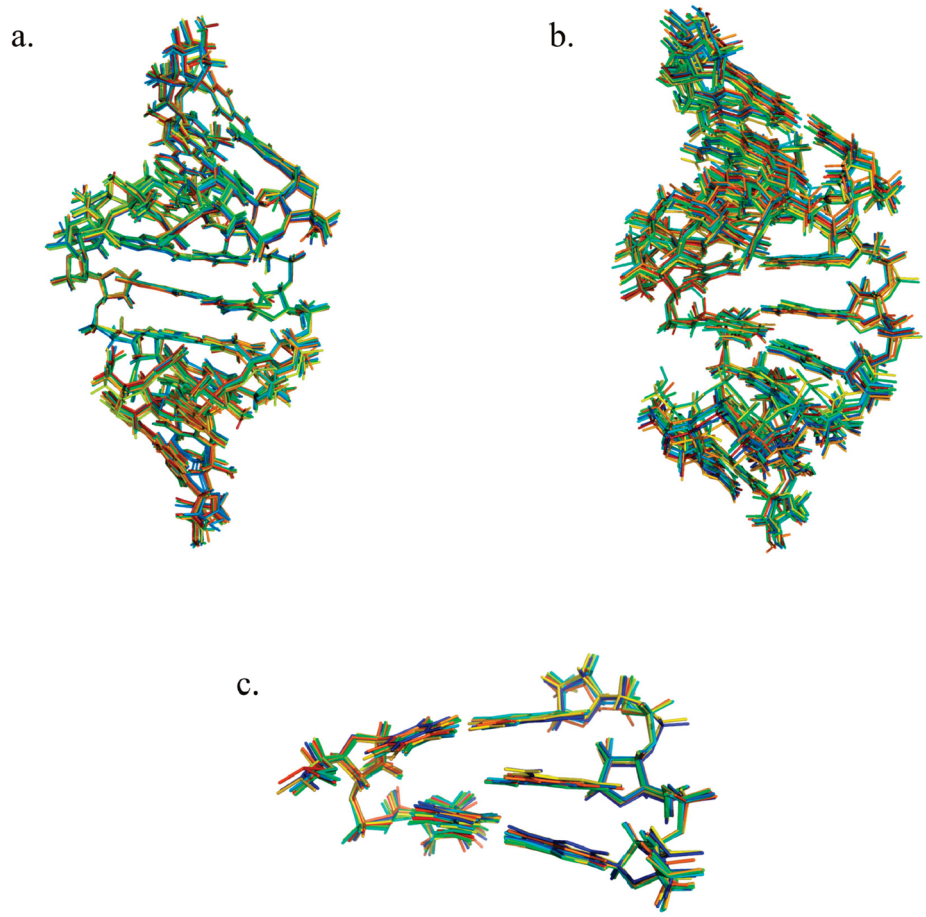

FIGURE 3: Superposition of 10 lowest energy structures of the reference (a) and bulged duplexes (b). Bulge region: superposition of 10 final structures. The view is into the major groove (c).

NOE connectivities in the upper strand is possible only if the bulged residue stacks into duplex. The presence of the interstrand NOE cross-peak between A5-H2/C18-H1′ and A5-H2 with A5-H1′ and A6-H1′ as well as the presence of sequential aromatic−aromatic cross-peaks between G4-A5-A6-G7 strongly suggests that the bulged adenine base A5 points inward of the duplex. In addition, A6-H2 shows NOE to C18-H1′, which supports the earlier observation that A6, not the A5 residue, is base paired with U17. This contact would not be possible for A5-U17 base pair. All of the phosphorus nuclei resonate within a range of approximately 1 ppm (−4.6 to −3.7 ppm), suggesting that the presence of a bulge residue does not significantly distort the conformation of the backbone A-form.

Structure Calculations. Table 1 summarizes the statistics on the distance and dihedral constraints used in the structure calculations for the reference and bulged duplexes. Both structures were calculated using the restrained molecular dynamics protocol described in Material and Methods. In order to minimize the potential uncertainty that could be imposed by using different calculation procedures, we have determined both structures applying identical protocols and parameter sets. The only difference between these calculations was in the NMR-derived experimental constraints. The calculations used a total of 367 distance constraints and 212 dihedral angle constraints for the reference duplex, and 337 distance constraints and 213 dihedral angle constraints for the bulged duplex. Out of 50 starting structures with completely random backbone dihedral angles, 40 and 36 structures converged to low NOE and dihedral angle violation energies for the reference and the bulged duplexes, respectively. The structures were classified as converged if they were consistent with the NMR data and maintained correct stereochemistry. All converged structures had no NOE constraints violated by more than 0.3 Å and no dihedral angle violations larger than 5°. The converged structures were analyzed, and for both molecules, all structures were found to fall in single structure families. A superposition of the 10 lowest-energy structures of both molecules is shown in Figure 3. The average pairwise rmsd calculated for all heavy atoms for these 10 structures is 0.31 (±0.07) Å and 0.62 (±0.26) Å, for the reference and the bulged duplex, respectively.

The structure of the reference duplex is a regular right-handed helix with all of the hallmarks of an A-type helix. These futures include the *anti* conformation of all residues with only small variations along the duplex and the average value of about −155°, Watson−Crick base pairing preserved in all base pairs, and C3′-*endo* sugar conformation with the average pseudorotation angles and puckering amplitudes of 15° and 40°, respectively. The regularity of the duplex is supported by the analysis of backbone torsion angles as these all lay in the standard range for an A-type duplex, that is (α to ζ), *gauche*-, *trans*, *gauche*+, *gauche*+, *trans*, and *gauche*-, with only slight variations along each of the strands. Additionally, we described the duplex structures by their helical parameters using the program CURVES. We calculated the helical twist, x-displacement, axial rise, and base-

Table 2: Selected Helical Parameters for the Reference and Bulged Duplexes[a]

base pairs		x-displacement dx (Å)		inclination η (°)		propeller twist ω (°)	
reference	bulged	reference	bulged	reference	bulged	reference	bulged
G1-C20	G1-C21	−4.7 (0.1)	−4.3 (0.2)	19 (1)	13 (1)	−12 (3)	−1 (3)
C2-G19	C2-G20	−4.7 (0.1)	−4.2 (0.2)	17 (2)	13 (2)	−11 (2)	−21 (2)
A3-U18	A3-U19	−4.8 (0.1)	−4.2 (0.3)	19 (1)	12 (1)	−6 (2)	−13 (2)
G4-C17	G4-C18	−4.6 (0.1)	−4.3 (0.1)	17 (2)	11 (2)	−4 (2)	−11 (2)
A6-U16	A6-U17	−4.8 (0.1)	−4.8 (0.2)	20 (1)	17 (2)	−11 (2)	−8 (3)
G7-C15	G7-C16	−4.4 (0.1)	−4.2 (0.2)	18 (2)	17 (2)	−6 (3)	−12 (2)
A8-U14	A8-U15	−4.8 (0.1)	−4.6 (0.1)	21 (2)	15 (2)	−9 (2)	−1 (2)
G9-C13	G9-C14	−4.4 (0.2)	−4.1 (0.2)	17 (2)	15 (2)	−7 (2)	−7 (3)
C10-G12	C10-G13	−4.3 (0.1)	−4.3 (0.2)	20 (2)	15 (2)	−25 (2)	−21 (3)
G11-C11	G11-C12	−4.6 (0.2)	−4.2 (0.2)	19 (1)	16 (2)	−1 (3)	−7 (4)
average:		−4.6 (0.2)	−4.3 (0.2)	19 (2)	15 (2)	−9 (7)	−10 (7)

base steps		rise D_z (Å)		twist Ω (°)		shift ρ (Å)	
reference	bulged	reference	bulged	reference	bulged	reference	bulged
G1-C2	G1-C2	2.6 (0.1)	2.6 (0.1)	36 (1)	40 (1)	0.0 (0.1)	0.2 (0.1)
C2-A3	C2-A3	2.8 (0.1)	2.8 (0.2)	30 (2)	29 (1)	−0.1 (0.1)	0.2 (0.1)
A3-G4	A3-G4	2.3 (0.1)	2.3 (0.1)	34 (2)	36 (2)	0.2 (0.2)	0.1 (0.3)
G4-A5		2.5 (0.1)		30 (2)		0.0 (0.1)	
A5-G6	A6-G7	2.6 (0.1)	2.6 (0.2)	34 (2)	29 (1)	0.4 (0.2)	0.4 (0.3)
G6-A7	G7-A8	2.6 (0.1)	2.1 (0.1)	32 (2)	34 (2)	−0.5 (0.2)	−0.1 (0.1)
A7-G8	A8-G9	2.4 (0.1)	2.6 (0.1)	30 (2)	29 (1)	0.6 (0.2)	0.9 (0.3)
G8-C9	G9-C10	2.6 (0.1)	2.9 (0.1)	36 (2)	35 (1)	0.3 (0.2)	−0.2 (0.2)
C9-G10	C10-G11	2.8 (0.2)	2.4 (0.1)	34 (1)	35 (1)	−0.6 (0.1)	0.0 (0.1)
average:		2.6 (0.2)	2.5 (0.2)	33 (3)	33 (4)	0.2 (0.4)	0.2 (0.4)

[a] Standard deviations from the family of structures calculated are given in parentheses.

pair inclination (Table 2). Base pair x-displacement of −4.6 ± 0.2 Å into the minor groove and helical twist angle 33 ± 3° were both characteristic of the A conformation. Also the axial rise (2.6 ± 0.2 Å) and base pair inclinations (19 ± 2°) classified the reference duplex as a typical A-structure.

Detailed analysis of the family of structures obtained for the bulged duplex reveals that the unpaired A5 adenosine residue is stacked into the helix in all converged structures. However, three out of 36 structures differ and show considerable deviations of the α and γ angles near the bulge site from the typical A-type values. For these three structures, the values of α (86°) and γ (−142°) angles in the adenosine residue A6 are exceptional. Although the lack of a downfield-shifted ^{31}P resonance in the bulged duplex suggested that in the bulge region none of the α and ζ-angles adopted the *trans* conformation, these angles were left unconstrained in the calculations. Additionally, the angles β and γ were also left unconstrained for A5 and A6 residues. Inspection of the remaining 33 structures shows that the γ angle of A6 residue consistently falls out of the normal *gauche*+ region with a mean value of ∼90°. The value of −113° for the α angle of A6 residue also differs from this typical of A-type structure. This result is a consequence of a dynamic nature of the sugar−phosphate backbone at the bulge site and of a greater conformational flexibility allowed for both α and γ backbone conformations. Comparison of the remaining torsion angles shows very good coherence between the reference duplex and the bulged structure.

In general, the helical parameters of the bulged duplex are very similar to those determined for the reference duplex and maintain analogous sequence-dependent variations throughout the helix. However, a detailed comparison of the parameters reveals some differences between these two structures. The accommodation of the bulge within the helix is manifested as a reduced inclination of all base pairs (average ∼ 4°). Additionally, although very slight but

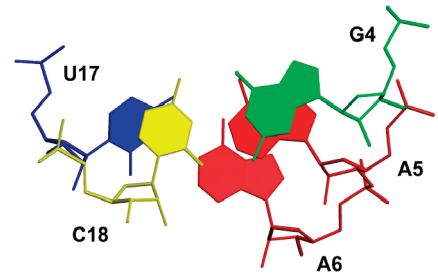

FIGURE 4: Stacking pattern of adjacent bases at the bulge site.

systematically reduced along the helix, base pair x-displacement (average ∼0.3 Å) is observed in comparison to that in the reference duplex. These two parameters illustrate the propagation of a local disturbance of the helix over the entire structure in order to minimize distortions induced by the incorporation of an unpaired residue into the regular helical structure.

Analysis of the interbase parameters provides additional information about the position of bulged adenine A5. The average shift for base−base steps is 0.2 ± 0.4 Å, whereas the shift of −1.9 ± 0.3 Å is observed at the A5/A6 step. Additionally, the twist angle at the A5/A6 step is considerably reduced form the average value of 32 ± 3° to 18 ± 4°. These local deviations of helical parameters observed at the bulge site arise from the unique arrangement of bases. The stacking of bases contributes to the stability of RNA as it minimizes the exposure of the hydrophobic base surface to the polar solvent. Intrahelical position of bulged adenine A5 is stabilized by more efficient stacking with the 5′-neighboring residue G4, then with the following residue A6. Stacking interactions among residues G4, A5, and A6 are presented in Figure 4.

The data from different biophysical methods including gel electrophoresis (*44*−*47*), fluorescence resonance energy transfer (FRET) measurements (*48, 49*), transient electric

birefringence (*50*), and NMR spectroscopy (*16*) show that the presence of bulges introduces a defined bend or kink into the helical axis. In our study, the adenosine bulge is found to induce only a small bend. Overall, bending angles are 19 ± 6°and 12 ± 2° for the bulged and the reference duplexes, respectively. However, we do not want to overestimate these results as we realize that the short-range nature of the NOE information does not allow a precise definition of the helix.

DISCUSSION

A growing number of experimentally determined crystal structures of RNAs, including ribosomal units, provide a wealth of information about the structure of RNA molecules. However, in comparison to X-ray structures, there are still a relatively small number of RNA structures obtained by NMR methods. As far as bulges are concerned, their loop-out conformation in a crystal might differ from that in a solution even for identical sequences. This is because a bulged base may be involved in a variety of interactions, for example, with proteins or metal ions and is influenced by crystal packing forces.

Conformation at the bulge site can depend on the type of the bulged residue, the identity of the nucleosides flanking the bulge, the sequence not adjacent to the bulge, and the length of the helix. In order to compare the conformation of the bulge established in this work to other single adenine bulges obtained in solution, we used a new RNA database called the RNA FRABASE (RNA FRAgments search engine and dataBASE; http://rnafrabase.ibch.poznan.pl) (*51*). All RNA structures that are deposited in the Protein Data Bank (*52*) were searched for the NMR structures with single adenine bulges flanked by two Watson−Crick or GU wobble base pairs. As a result, only six such motifs were identified (PDB IDs: 17RA (*19*), 1LMV (*22*), 1K8S (*21*), 1NC0 (*18*), 1D0U (*20*), and 1SLP (*17*)). All NMR structures, except for 1SLP, exhibit intrahelical orientation of the adenine bulge. In the 1SLP structure, the bulge is a part of a nine-nucleotide hairpin loop of a spliced leader RNA hairpin, and therefore, it is excluded from our discussion. The conformation and dynamics of the remaining single adenosine bulges differ depending on the types of flanking bases. When two GC or CG base pairs are adjacent to a bulged adenosine, 5′-GAG-3′/5′-CC-3′ or 5′-CAC-3′/5′-GG-3′ motifs (1K8S, 1NC0, and 1D0U), both imino proton resonances of the neighboring guanosines could be easily identified as involved in the hydrogen bond with the cytidine on the opposite strand. A different situation arises in the structures containing a distinctive structural motif: two adjacent adenine residues on one strand and a single uridine on the complementary strand. In principle, two secondary structures are possible because the bulge can be any of these two adenine residues. In one of the structures (1LMV), in the 5′-UAAC-3′/5′-GUA-3′ sequence context, a continuous A-type helical geometry was observed with both adenosines stacked within the helix. Both adenosine residues were found to form hydrogen bonds with the opposite uridine. Because no distinct second set of NOEs as evidence of conformational heterogeneity was observed in the NMR spectra, in regard to the paucity of NOEs involving exchangeable protons in the bulge region and the broadness of H2′, H3′, and H4′ proton resonances of both adenine residues, the authors (*22*) suggested the presence of some conformational flexibility at the bulge site. This dynamics was attributed to the possibility that the base flips in and out of the helix. In the 17RA (*19*) structure with 5′-UAAG-3′/5′-CUA-3′ motif, many of the exchangeable resonances in the bulge region were weak or not observable because of the broadening by exchange with water, including the guanosine imino resonance flanking the bulge from the 3′-side and uridine imino proton on the opposite strand. This prevented the direct identification of the AU base pair at the bulge site. Both adenine bases stacked into the helix, and while all available NOE and coupling data were consistent with both possible AU base pairs, the base pair involving the 5′-proximal adenine was suggested to be the major conformation. Because the 3′-proximal bulged adenine got protonated at an unusually high pH and could not be involved in the formation of a standard AU base pair, a model was proposed in which the protonated adenine was stabilized by a hydrogen bond to the uridine O2 of AU base pair.

In the structure containing the 5′-GAAG-3′/5′-CUC-3′ motif, studied in this work, we are able to unambiguously conclude that it is the A5 not the A6 adenine residue that remains unpaired, taking into consideration the NOE from uridine imino proton exclusively to adenine A6 H2. This finding is in agreement with the suggestion that single nucleotide bulges are preferentially located closer to the helix ends (*26*).

A comparison of the known solution structures, containing a single adenine bulge enclosed with two Watson−Crick or GU base pairs, provides some insight into the preferential conformation of the bulge site. When the bulged residue is not identical to either of the neighboring nucleotides, it forms stable, intrahelical structures with well-defined base pairs adjacent to the bulged adenine. If two neighboring adenine residues face one uridine in a complementary strand, a conformational diversity is possible: each of two adenine residues can base pair with uridine on the opposite strand, the uridine residue may form hydrogen bonds with two adenines, or there might be equilibrium between these conformations.

Following a suggestion of one of the reviewers, all solution structures discussed above have been searched for a possible correlation between the α and γ torsion angles around the bulge site. According to the earlier reported computational studies on single-base bulges (*53*), there is a correlation between the α and γ torsion angles. For two classes of bulges, i.e., when the base is stacked between flanking helices and when the base is contacting the minor groove, α and γ are found either in standard *gauche*− and *gauche*+ conformations, respectively, or both angles are close to the *trans* region. In the third class of structures with looped out bulges, the difference between α and γ was negative, and the α/γ flip was observed. The values of the α and γ torsion angles at the bulge site were obtained from the RNA FRABASE database (*51*). Their analysis shows only partial coherence with the data based on the computational studies of single-base bulges (*53*). In two structures (1K8S and 1D0U), the α and γ angles of both bulged and 3′ -neighboring residues show a tendency to adopt the *trans* conformation. In the 17RA structure, this propensity has been found only for the residue adjacent to the bulge from the 3′-site. No such relationship has been observed for the 1LMV structure, where

both α and γ angles fall in typical *gauche*− and *gauche*+ regions, respectively. In the two structures, our structure (2JXS) and 1NC0, the bulged adenosine, adopts a typical *gauche*−, *gauche*+ conformation, and the residue adjacent to the bulge from the 3′-site deviates from standard values, but with a tendency toward being closer to *gauche*−, *gauche*+ than *trans*.

In conclusion, a comparison of known solution structures with a single adenine bulge indicates that single adenine bulges can form structural hinges of different flexibility depending on the sequence context. It seems that if the AA/U motif is closed by two GC base pairs, the structure is still relatively well defined by NMR, as we were able to identify by closing the A6-U17 base pair. However, for 5′-UAAC-3′/5′GUA-3′ and 5′-UAAG-3′/5′-CUA-3′ motifs (*19, 22*), the bulge region is conformationally more flexible and difficult to determine in solution using NMR methods.

ACKNOWLEDGMENT

The generous support of the Poznan Supercomputing and Networking Center is acknowledged.

SUPPORTING INFORMATION AVAILABLE

Tables with ^{1}H, ^{13}C, and ^{31}P chemical shifts. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Egli, M. (2004) Nucleic acid crystallography: current progress. *Curr. Opin. Chem. Biol. 8*, 580–591.
2. Turner, D. H. (1992) Bulges in nucleic acids. *Curr. Opin. Struct. Biol. 2*, 334–337.
3. Gutell, R. R., Cannone, J. J., Shang, Z., Du, Y., and Serra, M. J. (2000) A story: unpaired adenosine bases in ribosomal RNAs. *J. Mol. Biol. 304*, 335–354.
4. Hermann, T., and Patel, D. J. (2000) RNA bulges as architectural and recognition motifs. *Structure 8*, R47−R54.
5. Beaucage, S. L., and Caruthers, M. H. (1981) Deoxynucleotide phosphoramidites -a new class of key intermediates for deoxypolynucleotide synthesis. *Tetrahedron Lett. 22* 1859–1862.
6. Vicens, Q., and Westhof, E. (2003) Crystal structure of Geneticin bound to a bacterial 16 S ribosomal RNA A site oligonucleotide. *J. Mol. Biol. 326*, 1175–1188.
7. Valegard, K., Murray, J. B., Stockley, P. G., Stonehouse, N. J., and Lijias, L. (1994) Crystal structure of an RNA bacteriophage coat protein-operator complex. *Nature 371*, 623–626.
8. Peattie, D. A., Douthwaite, S., Garret, R. A., and Noller, H. F. (1981) A "bulged" double helix in a RNA-protein contact site. *Proc. Natl. Acad. Sci. U.S.A. 78*, 7331–7335.
9. Sudarsanakumar, Ch., Xiong, M., and Sundaralingam, M. (2000) Crystal structure of an adenine bulge in the RNA chain of a DNA-RNA hybrid, d(CTCCTCTTC)-r(gaagagagag). *J. Mol. Biol. 299*, 103–112.
10. Portmann, S., Grimm, S., Workman, C., Usman, N., and Egli, M. (1996) Crystal structures of an A-form duplex with single-adenosine bulges and a conformational basis for site-specific RNA self-cleavage. *Chem. Biol. 3*, 173–184.
11. Ennifar, E., Yusupov, M., Walter, P., Marquet, R., Ehresmann, B., Ehresmann, C., and Dumas, P. (1999) The crystal structure of the dimerization initiation site of genomic HIV-1 RNA reveals an extended duplex with two adenine bulges. *Structure 7*, 1439–1449.
12. Joshua-Tor, L., Rabinovich, D., Hope, H., Frolow, F., Appela, E., and Sussman, J. L. (1988) The threedimensional structure of a DNA duplex containing looped-out bases. *Nature 334*, 82–84.
13. Cate, J. H., Gooding, A. R., Podell, E., Zhou, K. H., Golden, B. L., Kundrot, C. E., Cech, T. R., and Doudna, J. A. (1996) Crystal structure of a group I ribozyme domain: principles of RNA packing. *Science 273*, 1678–1685.
14. Rosen, M. A., Live, D., and Patel, D. J. (1992) Comparative NMR study of A(n)-bulge loops in DNA duplexes: intrahelical stacking of A, A-A, and A-A-A bulge loops. *Biochemistry 31*, 4004–4014.
15. Gollmick, F. A., Lorenz, M., Dornberger, U., von Langen, J., Diekmann, S., and Fritzsche, H. (2002) Solution structure of dAATAA and dAAUAA DNA bulges. *Nucleic Acids Res. 30*, 2669–2677.
16. Dornberger, U., Hillisch, A., Gollmick, F. A., Fritzsche, H., and Diekmann, S. (1999) Solution structure of a five-adenine bulge loop within a DNA duplex. *Biochemistry 38*, 12860–12868.
17. Greenbaum, N. L., Radhakrishnan, I., Patel, D. J., and Hirsh, D. (1996) Solution structure of the donor site of a trans-splicing RNA. *Structure 4*, 725–733.
18. Sashital, D. G., Allmann, A. M., Van Doren, S. R., and Butcher, S. E. (2003) Structural basis for a lethal mutation in U6 RNA. *Biochemistry 42*, 1470–1477.
19. Smith, J. S., and Nikonowicz, E. P. (1998) NMR structure and dynamics of an RNA motif common to the spliceosome branch-point helix and the RNA-binding site for phage GA coat protein. *Biochemistry 37*, 13486–13498.
20. Smith, J. S., and Nikonowicz, E. P. (2000) Phosphorothioate substitution can substantially alter RNA conformation. *Biochemistry 39*, 5642–5652.
21. Thiviyanathan, V., Guliaev, A. B., Leontis, N. B., and Gorenstein, D. G. (2000) Solution conformation of a bulged adenosine base in an RNA duplex by relaxation matrix refinement. *J. Mol. Biol. 300*, 1143–1154.
22. Newby, M. I., and Greenbaum, N. L. (2002) Sculpting of the spliceosomal branch site recognition motif by a conserved pseudouridine. *Nat. Struct. Biol. 9*, 958–965.
23. Barthel, A., and Zacharias, M. (2006) Conformational transitions in RNA single uridine and adenosine bulge structures: a molecular dynamics free energy simulation study. *Biophys. J. 90*, 2450–2462.
24. Hastings, W. A., Yingling, Y. G., Chirikjian, G. S., and Shapiro, B. A. (2006) Structural and dynamical classification of RNA single-base bulges for nanostructure design. *J. Comput. Theor. Nanosci. 3*, 63–77.
25. Merianos, H. J., Wang, J. M., and Moore, P. B. (2004) The structure of a ribosomal protein S8/spc operon mRNA complex. *RNA 10*, 954–964.
26. Serra, M. J., and Silvestri, S. B. (2002) Thermodynamic parameters for RNA bulge loop formation. *Biochemistry 41*, 91.
27. Longfellow, C. E., Kierzek, R., and Turner, D. H. (1990) Thermodynamic and spectroscopic study of bulge loops in oligoribonucleotides. *Biochemistry 29*, 278–285.
28. Zhu, J., and Wartell, R. M. (1999) The effect of base sequence on the stability of RNA and DNA single base bulges. *Biochemistry 38*, 15986–15993.
29. Usman, N., Ogilvie, K. K., Jiang, M.-Y., and Cedergren, R. J. (1987) Automated chemical synthesis of long oligoribonucleotides using 2′-O-silylated ribonucleoside 3′-O-phosphoramidites on a controlled-pore glass support: synthesis of a 43-nucleotide sequence similar to the 3′ half of an E. coli formylmethionine tRNA. *J. Am. Chem. Soc. 109*, 7845–7854.
30. Piotto, M., Saudek, V., and Sklenar, V. (1992) Gradient-tailored excitation for single-quantum NMR spectroscopy of aqueous solutions. *J. Biomol. NMR 2*, 661–665.
31. Shaka, A. J., Barker, P. B., and Freeman, R. (1985) Computer-optimized decoupling scheme for wideband applications and low-level operation. *J. Magn. Reson. 64*, 547–552.
32. Sklenar, V., Miyashiro, H., Zon, G., Miles, T., and Bax, A. (1986) Assignment of the 31P resonances in oligonucleotides by two-dimensional NMR Spectroscopy. *FEBS Lett. 208*, 94.
33. Clore, G. M., and Kuszewski, J. (2003) Improving the accuracy of NMR structures of RNA by means of conformational database potentials of mean force as assessed by complete dipolar coupling cross-validation. *J. Am. Chem. Soc. 125*, 1518–1525.
34. Varani, G., Jr. (1991) RNA structure and NMR spectroscopy. *Q. Rev. Biophys. 24*, 479–532.
35. Wijmenga, S. S., and van Buuren, B. N. M. (1998) The use of NMR methods for conformational studies of nucleic acids. *Prog. Nucl. Magn. Reson. Spectrosc. 32*, 287–387.
36. Roongta, V. A., Jones, C. R., and Gorenstein, D. G. (1990) Effect of distortions in the deoxyribose phosphate backbone conformation of duplex oligodeoxyribonucleotide dodecamers containing GT, GG, GA, AC, and GU base-pair mismatches on 31P NMR spectra. *Biochemistry 29*, 5245–5258.

37. Varani, G., Aboulela, F., and Allain, F. H. T. (1996) NMR investigation of RNA structure. *Prog. Nucl. Magn. Reson. Spectrosc. 29*, 51–127.
38. Stein, E. G., Rice, L. M., and Brunger, A. T. (1997) Torsion-angle molecular dynamics as a new efficient tool for NMR structure calculation. *J. Magn. Reson. 124*, 154–164.
39. Schwieters, C. D., Kuszewski, J. J., and Clore, G. M. (2006) Using Xplor-NIH for NMR molecular structure determination. *Prog. Nucl. Magn. Reson. Spectrosc. 48*, 47–62.
40. Schwieters, C. D., Kuszewski, J. J., Tjandra, N., and Clore, G. M. (2003) The Xplor-NIH NMR molecular structure determination package. *J. Magn. Reson. 160*, 65–73.
41. Parkinson, G., Vojtechovsky, J., Clowney, L., Brunger, A. T., and Berman, H. M. (1996) New parameters for the refinement of nucleic acid-containing structures. *Acta Crystallogr., Sect. D 52*, 57–64.
42. Lavery, R., and Sklenar, H. (1988) The definition of generalized helicoidal parameters and of axis curvature for irregular nucleic acids. *J. Biomol. Struct. Dyn. 6*, 63–91.
43. Lavery, R., and Sklenar, H. (1989) Defining the structure of irregular nucleic acids: conventions and principles. *J. Biomol. Struct. Dyn. 4*, 655–667.
44. Bhattacharyya, A., and Lilley, D. M. (1989) The contrasting structures of mismatched DNA sequences containing looped-out bases (bulges) and multiple mismatches (bubbles). *Nucleic Acids Res. 17*, 6821–6841.
45. Hsieh, C.-H., and Griffith, J. D. (1989) Deletions of bases in one strand of duplex DNA, in contrast to single-base mismatches, produce highly kinked molecules: possible relevance to the folding of single-stranded nucleic acids. *Proc. Natl. Acad. Sci. U.S.A. 86*, 4833–4837.
46. Luebke, K. J., Jr. (1996) Sequence effect on RNA bulge-induced helix bending and a conserved five-nucleotide bulge from the Group I Introns. *Biochemistry 35*, 11677–11684.
47. Wang, Y.-H., and Griffith, J. D. (1991) Effects of bulge composition and flanking sequence on the kinking of DNA by bulged bases. *Biochemistry 30*, 1358–1363.
48. Gohlke, C., Murchie, A. I., Lilley, D. M., and Clegg, R. M. (1994) Kinking of DNA and RNA helices by bulged nucleotides observed by fluorescence resonance energy transfer. *Proc. Natl. Acad. Sci. U.S.A. 91*, 11660–11664.
49. Lilley, D. M. (1995) Kinking of DNA and RNA by base bulges. *Proc. Natl. Acad. Sci. U.S.A. 92*, 7140–7142.
50. Zacharias, M., and Hagerman, P. J. (1995) Bulge-induced bends in RNA: quantification by transient electric birefringence. *J. Mol. Biol. 247*, 486–500.
51. Popenda, M., Blazewicz, M., Szachniuk, M., and Adamiak, R. W. (2008) RNA FRABASE version 1.0: an engine with a database to search for the three-dimensional fragments within RNA structures. *Nucleic Acids Res. 36*, D386–D391.
52. Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N., and Bourne, P. E. (2000) The Protein Data Bank. *Nucleic Acids Res. 28*, 235–242.
53. Zacharias, M., and Sklenar, H. (1999) Conformational analysis of single-base bulges in A-form DNA and RNA using hierarchical approach and energetic evaluation with a continuum solvent model. *J. Mol. Biol. 289*, 261–275.

BI7024904