Thermodynamic Stability and Structural Features of the J4/5 Loop in a *Pneumocystis carinii* Group I Intron†

Susan J. Schroeder,‡,§ Matthew A. Fountain,‖ Scott D. Kennedy,⊥ Peter J. Lukavsky,# Joseph D. Puglisi,△ Thomas R. Krugh,‡ and Douglas H. Turner*,‡

Department of Chemistry, University of Rochester, Rochester, New York 14627-0216, Department of Chemistry, State University of New York at Fredonia, Fredonia, New York 14063, Department of Biochemistry and Biophysics, University of Rochester, Rochester, New York 14642, MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB22QH, U.K., and Department of Structural Biology, Stanford University School of Medicine, Stanford, California 94305-5126

Received June 26, 2003; Revised Manuscript Received September 30, 2003

ABSTRACT: The J4/5 loop of the group I intron in the mouse-derived fungal pathogen *Pneumocystis carinii* is the docking site for the first step of the RNA-catalyzed self-splicing reaction and thus is a model of a potential drug target. This purine-rich asymmetric internal loop, 5′G**GAA**G/3′U**AG**U, is also thermodynamically more stable than other internal loops with two GU closing pairs and three nucleotides opposite two nucleotides. The results from optical melting, nuclear magnetic resonance spectroscopy, and functional group substitution experiments suggest that the GU closing pairs form and that sheared GA pairs form in the internal loop. The NMR spectra show evidence of conformational dynamics, and several GA pairings are possible. Thus, this dynamic loop presents several possible structures for potential binding of drugs that target group I self-splicing introns. The results also contribute to understanding the structural and dynamic basis for the function and thermodynamic stability of this loop.

The J4/5[1] loop in group I introns (Figure 1) is the docking site for the first step of the self-splicing reaction required for processing of ribosomal RNA (*1*−*6*). The structure and function of the J4/5 loop in the *Tetrahymena thermophila* group I intron has been extensively studied (*1*, *3*, *7*−*14*); this asymmetric internal loop (Figure 1, C2) contains tandem sheared adenine pairs. The J4/5 loop in the mouse-derived fungal pathogen *Pneumocystis carinii* (Figure 1, C1) has potential GA pairs and GU closing pairs and is thermodynamically more stable than the *T. thermophila* and *Candida albicans* J4/5 loops (*15*) (Figure 1, C2 and C3). The free energy of a tertiary contact to a GU pair at the group I intron

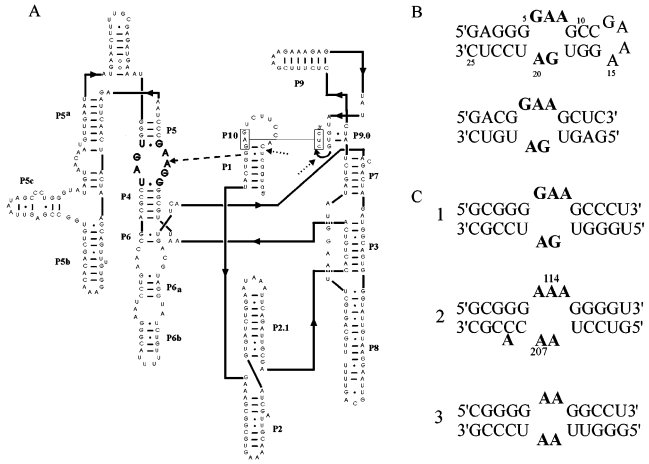

FIGURE 1: (A) Predicted secondary structure of the mouse-derived *P. carinii* group I intron in the large subunit rRNA precursor (*18*). The 5′ and 3′ exons are in lower case letters. The bold solid arrows show the direction of the sequence 5′ to 3′. The dotted arrows show the splice sites. The dashed arrow connects the P1 helix to the J4/5 loop; the P1 helix docks into the J4/5 loop in the first step of the self-splicing reaction. The nucleotides of the J4/5 loop are in large, bold letters. The helices are numbered according to standard convention (*1*). (B) Hairpin and non-self-complementary duplex models of the J4/5 loop in *P. carinii* used in these studies. (C) Comparison of the J4/5 loops in (1) *P. carinii* (*18*), (2) *T. thermophila* (*1*, *11*), and (3) *C. albicans* (*16*).

† This work supported by NIH Grant GM22939.

* To whom correspondence should be addressed. Tel: (585) 275-3207. Fax: (585) 506-0205. E-mail: Turner@chem.rochester.edu.

‡ Department of Chemistry, University of Rochester.

§ Current address: Department of Chemistry, Yale University, P.O. Box 208107, New Haven, CT 06520-8107.

‖ State University of New York at Fredonia.

⊥ Department of Biochemistry and Biophysics, University of Rochester.

MRC Laboratory of Molecular Biology.

△ Stanford University School of Medicine.

[1] Abbreviations: COSY, correlation spectroscopy; DQFCOSY, double-quantum-filtered correlation spectroscopy; DTT, dithiothreitol; EDTA, ethylenediaminetetraacetic acid; eu, entropic units (cal K^{-1} mol^{-1}); HETCOR, heteronuclear correlation; HPLC, high-pressure liquid chromatography; HSQC, heteronuclear single-quantum coherence; *m*X*n* loop, an internal loop containing *m* nucleotides opposite *n* nucleotides, where *m* and *n* are integers; J4/5, junction between helices P4 and P5 in a group I intron; MWCO, molecular weight cutoff; NAIM, nucleotide analogue interference mapping; NAIS, nucleotide analogue interference suppression; NMR, nuclear magnetic resonance spectroscopy; NOE, nuclear Overhauser effect; NOESY, nuclear Overhauser spectroscopy; NTP, nucleotide triphosphates; PEG, poly(ethylene glycol); ROESY, rotating frame nuclear Overhauser spectroscopy; RMSD, root mean square deviation; SNOESY, nuclear Overhauser spectroscopy with an S-shaped pulse for water suppression; TLC, thin-layer chromatography; T_M, melting temperature in kelvin; T_m, melting temperature in degrees Celcius; TOCSY, total correlation spectroscopy.

10.1021/bi0301587 CCC: $25.00 © 2003 American Chemical Society
Published on Web 11/13/2003

splice site (*2*, *16*, *17*) correlates with the J4/5 loop stabilities (*15*). Thus, structural studies of the J4/5 loop of *P. carinii* can provide insight for understanding sequence−structure−energetics−function relationships in internal loops.

Internal loops have many biological functions, including binding proteins and forming tertiary interactions that determine global folding of RNA. Thus, understanding the structure and dynamics of internal loops can contribute to the rational choice of internal loops for drug target sites and the design of drugs to bind internal loops and inhibit the biological function of the loop. *P. carinii* is an opportunistic pathogen that infects immunocompromised patients. The precursor rRNA in *P. carinii* contains a group I intron, and inhibition of the group I intron self-splicing reaction is a strategy for drug design (*16*−*26*).

The J4/5 loop contains several phylogenetically conserved nucleotides (*1*, *11*, *27*). In an analysis of 71 group I introns, 100% contain purine-rich internal loops at this site. Moreover, 97% of J4/5 loops conserve adenines at positions 114 and 207 (*10*) (*T. thermophila* numbering; see Figure 1, C2). A model of the J4/5 loop in *T. thermophila* based on NAIM (nucleotide analogue interference mapping) and NAIS (nucleotide analogue interference suppression) results proposes a role for A114 and A207 as hydrogen bond acceptors in docking of the splice site GU pair and for A207 in stabilizing the transition state (*3*, *9*, *28*). When guanine is present, it occurs most frequently at positions 113 and 206, as found in the *P. carinii* J4/5 loop. Of 71 J4/5 loops, 80% contain at least one GU closing pair, and 80% contain three or more consecutive G bases adjacent to the loop on either the 5′ or 3′ side. NAIM results also suggest an important role for the amino groups in the guanines adjacent to the internal loop in *T. thermophila* (*29*, *30*).

To investigate the structure and thermodynamics of an isolated J4/5 loop, a hairpin and a duplex model were designed that contain the J4/5 loop from mouse-derived *P. carinii* (Figure 1, B). The NMR spectroscopy, functional group substitution, and optical melting studies reported here address the following questions: (1) What is the basis for the thermodynamic stability of the internal loop? (2) What are the conformations of the base pairs within the internal loop? (3) Are important functional groups presented in an arrangement for docking similar to that in the *T. thermophila* J4/5 loop?

The NMR spectra suggest the internal loop is conformationally dynamic but with some well-defined structural features. One possible conformation of the loop contains sheared GA pairs similar to the tandem sheared AA pairs in the *T. thermophila* J4/5 loop. This hypothesis was tested with 7-deazaadenine and inosine substitutions for the adenine and guanine nucleotides, respectively, within the loop. Removal of the N7 hydrogen bond acceptor by 7-deazaadenine substitution or of the amino group hydrogen bond donor by inosine substitution should disrupt formation of sheared GA pairs (Figure 2). These substitutions in the J4/5 loop in *P. carinii* destabilize the loop thermodynamically and change the structure and dynamics of the loop. The formation of sheared GA pairs adjacent to GU closing pairs and the stacking of the loop nucleotides on the adjacent helices contribute to the thermodynamic stability of the loop. Thus, models of a dynamic J4/5 loop containing possible sheared GA pairs are presented.

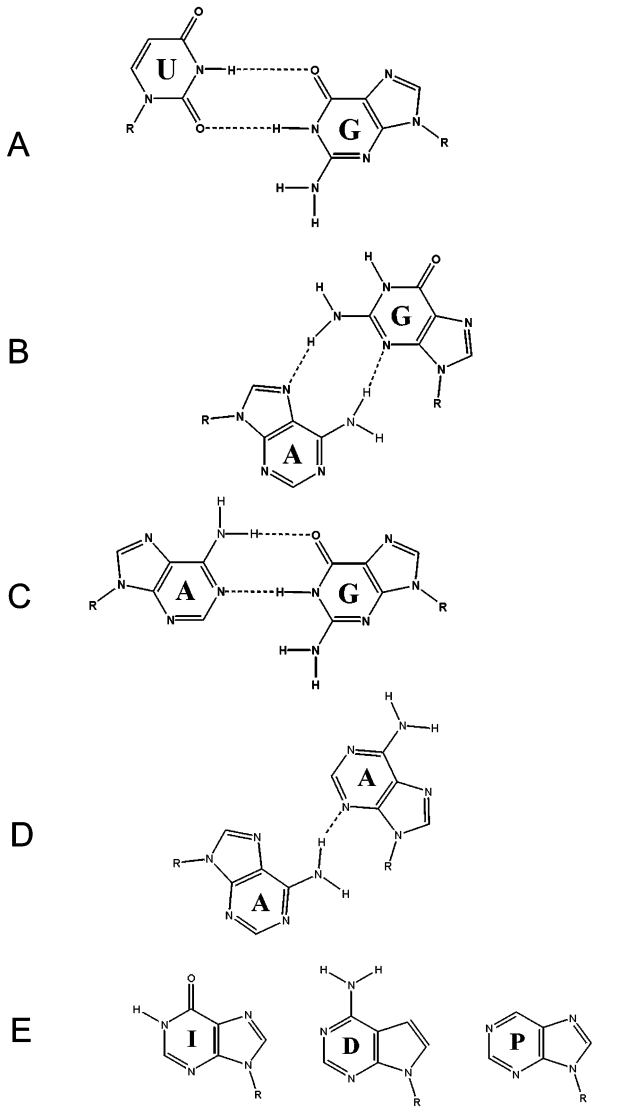

FIGURE 2: (A) Wobble GU pair, (B) sheared GA pair (N7-amino, amino-N3, trans Hoogsteen/sugar edge) (*77*, *89*, *90*), (C) imino GA pair (N1-N1, carbonyl-amino, cis Watson−Crick/Watson−Crick) (*77*, *89*, *90*), (D) sheared AA pair as seen in the crystal structure of the J4/5 loop in *T. thermophila* (*8*), and (E) inosine (I), 7-deazaadenine (D), and purine (P).

MATERIALS AND METHODS

Oligoribonucleotide Synthesis. Oligoribonucleotides were synthesized on an ABI 392 DNA/RNA synthesizer using phosphoramidite chemistry (*31*, *32*). Incubation in 3:1 (v/v) ammonia−ethanol at 55 °C overnight removed the base-protecting groups and the CPG support; incubation in triethylamine−hydrogen fluoride at 55 °C for 48 h removed the silyl-protecting group on the 2′-hydroxyls. Oligomers were desalted on a Sep-Pak C-18 cartridge and purified on a Baker 500Si TLC plate using 55:35:10 1-propanol−ammonia−water solvents. The purity of 10- and 11-mers was analyzed by HPLC; the purity of 25-mers was checked by ^{32}P labeling and gel electrophoresis. All oligomers were greater than 90% pure. The inosine and 7-deazaadenine phosphoramidites were purchased from Glen Research Corp. and ChemGenes Corp., respectively, and deblocked with the standard protocol. Four 1 μmol syntheses were sufficient for an approximately 0.8 mM NMR sample (250 μL) of 25-mer oligomers with inosine or 7-deazaadenine substitutions.

For NMR samples of 25-mer oligonucleotides with unmodified bases, 20−25 mL T7 transcription reactions yielded approximately 0.9−1.8 mM NMR samples in 250 μL. Typical T7 transcription reaction conditions were the following (*33*, *34*): 16 mM nucleotide triphosphates, pH 8.1 (4 mM each NTP), 16 or 32 mM $MgCl_2$ (1:1 or 2:1 Mg^{2+}−NTP), 500 nM double-stranded DNA template with a penultimate 2′-*O*-methyl (*35*), approximately 80000 units/mL T7 enzyme, 20 mM DTT (dithiothreitol), 40 mM Tris, pH 8.1, 1 mM spermidine, 0.1% Triton X, 80 mg/mL PEG [poly(ethylene glycol)], and incubation at 37 °C for 1 h. The optimal amount of T7, DTT, and incubation time varies with each batch of T7 enzyme. A fully $^{13}C-^{15}N$-labeled NMR sample was prepared using NTPs extracted from *Methylophilus methylotrophus* (*36*−*38*); an adenine $^{13}C-^{15}N$-labeled NMR sample was prepared with 45 mg of $^{13}C-^{15}N$-labeled adenine triphosphates from Martec Biosciences Inc.

After the T7 transcription reaction, the RNA product was phenol extracted and desalted on a Sephadex column. The RNA was purified by gel electrophoresis for 12−15 h at 50−65 W on a denaturing 20% polyacrylamide gel (42.7 cm × 34.5 cm × 2 mm). The RNA was eluted from the gel in a Schleicher and Schuel Elutrap device. The RNA was desalted on a Sephadex PD10 column and dialyzed against 2 L of H_2O for 24 h and then 2 L of 40 mM NaCl, 5 mM phosphate, and 0.05 mM Na_2EDTA for 24 h with a microdialysis continuous-flow system and a 1000 MWCO dialysis membrane. The purity was confirmed to be greater than 95% by removal of the 5′-triphosphate with calf intestinal phosphatase, 5′-^{32}P labeling with T4 polynucleotide kinase, and gel electrophoresis.

Optical Melting Experiments. Concentrations of single strand oligonucleotides were determined from the absorbance at 280 nm at 80 °C (*39*). The extinction coefficients for 7-deazaadenosine and inosine were assumed to be those of adenosine and guanosine, respectively. Standard melt buffer was 1 M NaCl, 20 mM sodium cacodylate, and 0.5 mM Na_2-EDTA, pH 7. Melting curves were measured at 280 nm on a Gilford 250 spectrophotometer with a heating rate of 1 °C/min controlled by a Gilford 2527 thermoprogrammer.

Melting curves were fit to the two-state model with sloping baselines (*40*). Thermodynamic parameters were also obtained by fitting plots of inverse melting temperature, T_M^{-1}, versus natural log(C_T/n) to the equation (*41*):

$$T_M^{-1} = (R/\Delta H°) \ln(C_T/n) + \Delta S°/\Delta H° \qquad (1)$$

where R is the gas constant, C_T is total strand concentration, and n is 4 for non-self-complementary duplexes and 1 for self-complementary duplexes. Data from the melting experiments were analyzed with the program Meltwin 3.0 (*42*, *43*).

Gel Filtration. The 25-nucleotide RNA oligomer at approximately 1.8×10^{-4} M in 80 mM NaCl, 10 mM phosphate, and 0.5 mM Na_2EDTA, pH 6 (NMR buffer), was run on a Bio-Sil-125 gel filtration column from Bio-Rad Corp. using a Dynamax HPLC system. The running buffer was 50 mM sodium phosphate, 150 mM NaCl, and 1 mM EDTA; the flow rate was 1 mL/min. The RNA was prepared (1) at room temperature or (2) heated to 90 °C and snap cooled on ice or (3) heated to 90 °C and slow cooled for approximately 1 h after the addition of 1 M NaCl.

Nuclear Magnetic Resonance (NMR) Experiments. Short oligonucleotide duplex models of the J4/5 loop and loops containing 7-deazaadenine or inosine substitutions were studied by one-dimensional NMR spectroscopy with either an S-shaped pulse (*44*) or a binomial pulse sequence (*45*) for water suppression. A hairpin model of the J4/5 loop from *P. carinii* was studied with standard homonuclear and heteronuclear NMR spectroscopy methods (*46*−*49*). The detailed methods for the NMR spectroscopy experiments are available in Supporting Information. Hairpins containing 7-deazaadenine or inosine substitutions were studied with one- and two-dimensional homonuclear NMR spectroscopy experiments and natural abundance ^{13}C NMR spectroscopy experiments.

Modeling. Distance restraints for modeling were generated from peak lists for D_2O NOESY experiments at 50, 100, 150, 200, and 400 ms mixing times, SNOESY experiments at 100 and 150 ms mixing times, and ^{13}C-edited NOESY experiments at 100 and 400 ms mixing times. NOE intensities were estimated as representative of 1.5−3.0, 2.5−4.0, or 3.5−5.5 Å distances if the peak appears first at 50, 100, or 150 ms mixing times, respectively, in D_2O NOESY and ^{13}C-edited NOESY experiments (*47*). For SNOESY experiments, NOE intensities were approximated as 1.5−4.0 or 3.0−5.5 Å if the peak appears first at 100 or 150 ms mixing times, respectively. These distance approximations matched distances in A-form RNA for most NOEs of the Watson−Crick stem nucleotides. Some approximated distances were expanded to include the distance in A-form RNA, thus allowing for the possibility that relaxation effects or internal dynamics attenuated the peak intensity. Initial modeling used a total of 315 NOE distance restraints including restraints from weak NOEs observed at 200 or 400 ms mixing times. Only 14 weak restraints from NOEs observed at 200 or 400 ms mixing times were included in the final modeling, giving a total of 228 NOE distance restraints for the 25-mer RNA. Four of these are weak internucleotide restraints for nucleotides in the internal loop (refer to Supporting Information).

Loose dihedral angle restraints, α (−150 to 150 °), β (−150 to 150 °), γ (30 to 90 °), δ (70 to 110 °), ϵ (−170 to −100 °), ζ (−150 to 150 °), and χ (−150 to 150 °), were applied to the Watson−Crick stems. The presence of strong H3′−P_{n+1} cross-peaks in the ^{31}P HETCOR spectrum between −0.5 and −2.0 ppm (Supporting Information), as typically observed for A-form RNA, the lack of observable H1′−H2′ couplings in TOCSY and DQFCOSY spectra (Supporting Information), and weak or absent H1′−H8 cross-peaks when the H1′ resonance has strong intensity and strong NOEs to other resonances support the inclusion of the listed dihedral restraints (*50*). For internal loop nucleotides, including the GU closing pairs, loose dihedral angle restraints were applied only when supported by positive experimental evidence. Overlap in the ribose region precludes accurate measuring of coupling constants. No dihedral angle restraints were applied to the nucleotides in the hairpin loop. A total of 95 loose dihedral angle restraints were used to model the 25-mer RNA.

Hydrogen bond restraints of 1.6−2.0 Å were used for Watson−Crick pairs, GU pairs, and noncanonical pairs in the loop. The results of an HNNCOSY experiment (Supporting Information) and the patterns of NOEs in the helix stem regions support the inclusion of hydrogen bond

Table 1: Thermodynamic Parameters of Duplex Formation[a]

		T_M^{-1} vs ln(C_T/4) parameters			T_m (°C)	curve fit parameters			T_m (°C)
duplex	name	$-\Delta G°_{37}$ (kcal/mol)	$-\Delta H°$ (kcal/mol)	$-\Delta S°$ (eu)	(1.0 × 10^{-4} M)	$-\Delta G°_{37}$ (kcal/mol)	$-\Delta H°$ (kcal/mol)	$-\Delta S°$ (eu)	(1.0 × 10^{-4} M)
J4/5 P.c. loop									
5′GAGU **GA** UGAC[b] 3′CUCG**AAG**GCUG	J4/5 duplex	7.0 ± 0.1	75.6 ± 2.8	221.2 ± 9.2	38.8	7.0 ± 0.1	75.3 ± 7.5	220.0 ± 24.2	39.1
inosine substitutions									
5′GAGU **IA** UGAC 3′CUCG**AAG**GCUG	G5I duplex	6.0 ± 0.1	76.5 ± 3.9	227.3 ± 12.7	35.0	6.2 ± 0.2	66.8 ± 5.7	195.3 ± 19.2	35.6
5′GAGU **GA** UGAC 3′CUCG**AAI**GCUG	G15I duplex	5.8 ± 0.2	63.2 ± 6.1	185.2 ± 19.8	33.3	5.9 ± 0.2	60.4 ± 6.7	175.8 ± 22.1	33.8
5′GAGU **IA** UGAC 3′CUCG**AAI**GCUG		5.1 ± 0.2	55.0 ± 3.5	160.8 ± 11.6	29.2	5.3 ± 0.3	53.3 ± 9.3	154.7 ± 30.7	30.0
7-deazaadenine substitutions									
5′GAGU **GA** UGAC 3′CUCG**DAG**GCUG	A17D duplex	6.4 ± 0.3	86.5 ± 3.0	258.3 ± 9.8	36.6	6.5 ± 0.2	79.1 ± 6.2	234.0 ± 20.3	37.0
5′GAGU **GA** UGAC 3′CUCG**ADG**GCUG	A16D duplex	6.1 ± 0.1	85.7 ± 5.7	256.6 ± 18.7	35.5	6.3 ± 0.2	75.9 ± 7.2	224.5 ± 24.0	36.0
5′GAGU **GD** UGAC 3′CUCG**AAG**GCUG	A6D duplex	5.3 ± 0.1	66.7 ± 4.4	197.9 ± 14.5	31.4	5.6 ± 0.4	60.6 ± 9.4	177.6 ± 31.6	32.1
2 × 3 loop with GC closure 5′CGAC **GA** GCAG 3′GCUG**AAG**CGUC		9.7 ± 0.1	93.6 ± 1.6	270.3 ± 4.9	47.9	9.4 ± 0.1	79.9 ± 2.8	227.6 ± 9.2	48.2
2 × 2 loop 5′GCC**GD**GGC 3′CGG**DG**CCG		7.0 ± 0.1	64.9 ± 3.0	186.7 ± 9.6	43.6	6.9 ± 0.1	55.3 ± 6.1	155.9 ± 19.5	44.0

[a] Melting experiments were done in 1 M NaCl, 20 mM sodium cacodylate, and 0.5 mM Na_2EDTA, pH 7, buffer solution. Sequences are listed in order of type of loop and substitution and then by duplex free energy. "I" indicates an inosine substitution, and "D" indicates a 7-deazaadenine substitution. Listed errors are standard deviations from measurements assuming no correlation of errors in the slope and intercept and are therefore overestimates of this source of error. Estimated errors from all sources are ±10%, ±10%, ±2%, and ±1 °C for $\Delta H°$, $\Delta S°$, $\Delta G°$, and T_m, respectively. Significant figures beyond error estimates are given to allow accurate calculation of T_m and other parameters. The T_m listed is for a total strand concentration of 1.0×10^{-4} M. [b] Reference *15*. The buffer solution was 10 mM sodium cacodylate, 1 M NaCl, and 0.5 mM Na_2EDTA, pH 7.

restraints for the Watson−Crick and GU pairs, which helps to keep the base pairs planar. Several possible sets of hydrogen bond restraints were tested for the internal loop nucleotides because the results of nucleotide substitutions support the hypothesis that sheared GA pairs form within the internal loop. Hydrogen bond restraints were applied for N7-amino, amino-N3 (sheared) GA pairs to nucleotides G6−A20, and G19−A7 or G19−A8. Hydrogen bond restraints were also tested for either N7-amino, symmetric or N1-amino, amino-N7 AA pairs to nucleotides A7−A20. A total of 25 hydrogen bond restraints were used to model the 25-mer RNA in the final models.

The Insight2000 package containing NMRchitect, Discover, X-PLOR, Biosym, and AMBER95 on a Silicon Graphics Octane workstation was used for simulated annealing and analysis of the resulting structures. An initial structure was built and energy minimized with Biosym's Biopolymer module. To randomize the structure before running the simulated annealing protocol, a total of 0.25 ps of unrestrained dynamics was performed as the temperature increased from 200 to 1000 K in 20 equal increments with 1 fs steps while the Rattle algorithm (*51*) was applied. The simulated annealing protocol followed the procedure described by McDowell et al. (*42*); a detailed description of the simulated annealing protocol is included in Supporting Information. The nonbonded cutoff term was 12 Å, and the phosphate charges were those in AMBER95 (*52*). A flat-bottom square-well potential was used with force constants of 25 kcal/(mol·Å^2) and 25 kcal/(mol·rad^2) for distance and dihedral angle restraints with a maximum force of 1000 kcal/mol. The final minimization step used electrostatic terms and a Lennard-Jones nonbonding potential. A total of 50 structures with G6−A20 and A8−G19 hydrogen bond restraints, 50 structures with G6−A20 and A7−G19 hydrogen bond restraints, and 25 structures with no hydrogen bond restraints for GA pairs were generated for the final modeling.

RESULTS

Thermodynamics. Tables 1 and 2 list the thermodynamic parameters for duplex and loop formation, respectively, for a duplex model of the J4/5 loop from *P. carinii* and for related sequences. The thermodynamic parameters from T_M^{-1} vs ln(C_T/4) plots and from curve fits of the data agree within 15%, which is consistent with the two-state model assumption (*53*−*55*). The error in $\Delta G°_{37}$ for the nearest neighbor pair 5′GG/3′UU is ±1 kcal/mol (*56*), which accounts for the large error in the free energy increments of the internal loops. The error for the free energy for duplex formation is less; thus, the relative free energies of loop formation are more accurate when the stem sequences are the same. This relative error is given in parentheses in Table 2.

Figure 2 shows wobble GU, sheared GA, imino GA, and sheared AA pairs and inosine, 7-deazaadenine, and purine bases. Substitution of guanine by inosine in a sheared GA pair removes a hydrogen bond donor, while substitution of

Table 2: Thermodynamic Parameters for Internal Loop Formation[a]

duplex	name	$\Delta G^\circ_{loop,37}$ (kcal/mol)	ΔH°_{loop} (kcal/mol)	ΔS°_{loop} (eu)
J4/5 P.c. loop 5′GAGU **GA** UGAC[b] 3′CUCG**AAG**GCUG	J4/5 duplex	1.9 ± 1.0 (±0.1)	−14.2 ± 11.6 (±2.8)	−51.9 ± 27.9 (±9.2)
inosine substitutions 5′GAGU **IA** UGAC 3′CUCG**AAG**GCUG	G5I duplex	2.8 ± 1.0 (± 0.1)	−15.2 ± 11.9 (±3.9)	−58.0 ± 29.2 (±12.7)
5′GAGU **GA** UGAC 3′CUCG**AAI**GCUG	G15I duplex	3.1 ± 1.0 (±0.2)	−1.9 ± 12.8 (±6.1)	−15.9 ± 33.0 (±19.8)
5′GAGU **IA** UGAC 3′CUCG**AAI**GCUG		3.8 ± 1.0 (±0.2)	6.3 ± 11.8 (±3.5)	8.5 ± 28.8 (±11.6)
7-deazaadenine substitutions 5′GAGU **GA** UGAC 3′CUCG**DAG**GCUG	A17D duplex	2.5 ± 1.0 (±0.3)	−25.2 ± 11.6 (±3.0)	−89.0 ± 28.1 (±9.8)
5′GAGU **GA** UGAC 3′CUCG**ADG**GCUG	A16D duplex	2.8 ± 1.0 (±0.1)	−24.4 ± 12.6 (±5.7)	−87.3 ± 32.3 (±18.7)
5′GAGU **GD** UGAC 3′CUCG**AAG**GCUG	A6D duplex	3.6 ± 1.0 (±0.1)	−5.4 ±12.1 (±4.4)	28.6 ± 30.1 (±14.5)
2 × 3 loops with GC closure 5′CGAC **GA** GCAG 3′GCUG**AAG**CGUC		0.2 ± 0.2	−33.7 ± 2.8	−109.5 ± 8.4
5′GAGC **GA** CGAC[c] 3′CUCG**AAG**GCUG		0.2 ± 0.2	−19.78 ± 3.12	−64.38 ± 9.52
2 × 2 loops 5′GC**IA**GCG[d] 3′GCG**AI**CG		−2.2 ± 0.3	−24.8 ± 3.3	−73.1 ± 10.0
5′GC**GA**GCG[d] 3′GCG**AG**CG		−1.4 ± 0.2	−18.8 ± 4.9	−56.2 ± 8.2
5′GGC**GA**GCC[d] 3′CCG**AG**CGG		−0.7 ± 0.3	−8.9 ± 3.0	−26.5 ± 8.9
5′GCC**GD**GGC 3′CGG**DG**CCG		1.8 ± 0.5	−12.8 ± 8.3	−47.4 ± 24.9
5′GGC**GP**GCC[d] 3′CCG**PG**CGG		2.3 ± 0.1	0.4 ± 2.1	−6.0 ± 6.2

[a] Melting experiments were done in 1 M NaCl, 20 mM sodium cacodylate, and 0.5 mM EDTA, pH 7, buffer solutions. Sequences are listed in order of internal loop free energy for each type of loop and substitution. "I" indicates an inosine substitution, "D" indicates a 7-deazaadenine substitution, and "P" indicates a purine substitution. The experimental loop free energy is calculated according to the equation (*87*): $\Delta G^\circ_{loop} = \Delta G^\circ_{duplex\ with\ loop} - (\Delta G^\circ_{duplex\ without\ loop} - \Delta G^\circ_{interrupted\ base\ pair})$. Thermodynamic values used in this equation are from T_M^{-1} vs $\ln(C_T/n)$ plots. Estimated errors for the thermodynamic parameters of the interrupted nearest neighbor pair 5′CG/3′GC are ±1.7 kcal/mol, ±5 eu, and ±0.1 kcal/mol for ΔH°, ΔS°, and ΔG°_{37}, respectively. The errors in the nearest neighbor pair 5′GG/3′UU are ±8.4 kcal/mol, ±25.7 eu, and ±1 kcal/mol for ΔH°, ΔS°, and ΔG°_{37}, respectively. The error in the measurement of thermodynamic parameters for duplex formation (Table 1) is listed in parentheses; this value provides a better estimate of relative error due to the substitution in the loop and is independent of the error contributed to the loop parameters from the 5′GG/3′UU nearest neighbor pair, which is constant for all GU closed 2 × 3 loops in this table. [b] Reference *15*. Buffer solution was 10 mM sodium cacodylate, 1 M NaCl, and 0.5 mM Na_2EDTA, pH 7. [c] Reference *79*. [d] These values are from ref *57* and are recalculated with updated nearest neighbor parameters (*88*).

adenine by 7-deazaadenine removes a hydrogen bond acceptor. For both cases, the loss of a hydrogen bond is predicted to be thermodynamically destabilizing. In contrast, substitution of guanine by inosine in an imino GA pair is thermodynamically stabilizing, possibly because the pK_a of inosine (8.8) is more favorable for a hydrogen-bonding donor than the pK_a of guanine (9.2) (*57*). Similarly, substitution of adenine by 7-deazaadenine is predicted to be slightly stabilizing for an imino GA pair because the pK_a of 7-deazaadenine (5.3) is more favorable for a hydrogen-bonding acceptor than the pK_a of adenosine (3.5) (*58*). The inosine and 7-deazaadenine substitutions in the 2 × 3 loops are thermodynamically destabilizing (Table 2), which is more consistent with the formation of sheared GA pairs than imino GA pairs.

The change in free energy with the substitution of an inosine for a guanine or 7-deazaadenine for an adenine in sheared GA pairs results from a change in hydrogen bonding and/or conformation. The average change in free energy per inosine substitution in the 2 × 3 loop, 5′U**GA**U/3′G**AAG**G, is 1.0 kcal/mol at 37 °C, which is consistent with measurements of the free energy contribution from a hydrogen bond (*57*, *59*, *60*). The free energy change for an inosine substitution in a sheared GA pair in a GCAA hairpin is 0.5 kcal/mol at 37 °C (*59*). Inosine substitutions in the tandem sheared GA pairs of $(5'GC\mathbf{GA}GCG)_2$ enhance stability but are associated with a conformational change (*61*), probably to an imino IA conformation as observed in the crystal structure of 5′C**AI**G/3′G**IA**C (*62*).

The average change in free energy for each 7-deazaadenine substitution is 1.1 kcal/mol at 37 °C, which is consistent with the loss of a hydrogen bond. In the 5′C**GA**G/3′G**AG**C loop, the average change in free energy for each 7-deazaadenine or purine substitution is 1.2 and 1.5 kcal/mol at 37 °C, respectively (Table 2), which is consistent with the predicted contribution from a hydrogen bond (*57*, *59*, *60*). In the 2 ×

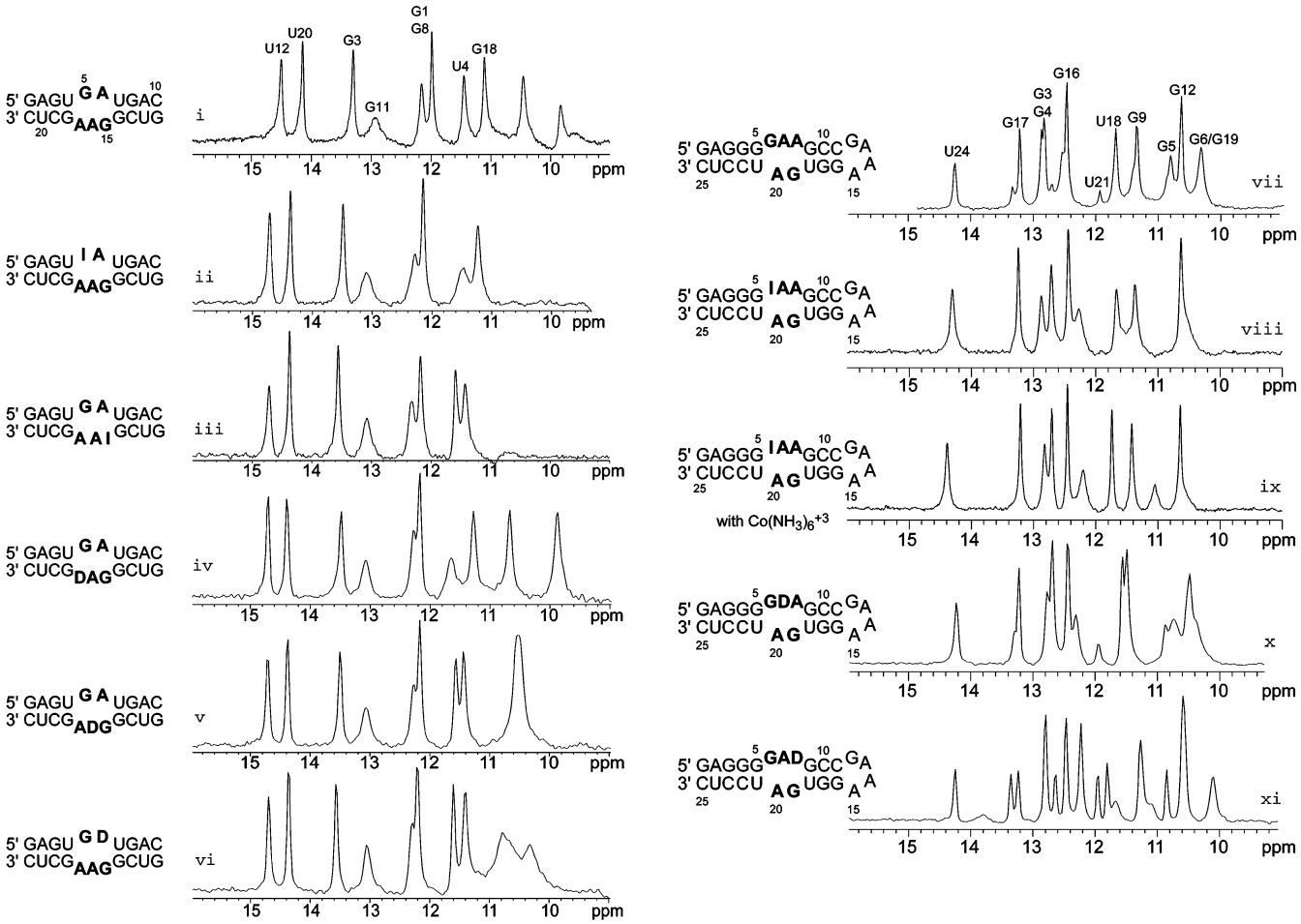

FIGURE 3: One-dimensional imino ^{1}H NMR spectra (9−16 ppm) of duplex and hairpin models of the J4/5 loop from *P. carinii* and loops containing nucleotide substitutions in 80 mM NaCl, 10 mM phosphate, and 0.5 mM Na_2EDTA, pH 6.0 buffer, unless otherwise noted. The single spectrum shown from a variable temperature study was selected for the best dispersion: (i) 2 mM J4/5 duplex, 5 °C and pH 6.7; (ii) 0.5 mM G5I duplex, 10 °C; (iii) 0.4 mM G15I duplex, 10 °C; (iv) 0.5 mM A17D duplex, 10 °C; (v) 0.5 mM A16D duplex, 10 °C; (vi) 0.5 mM A6D duplex, 10 °C; (vii) 1.5 mM J4/5 hairpin, 15 °C; (viii) 0.8 mM G6I hairpin, 10 °C; (ix) 0.8 mM G6I hairpin, 15 °C, in the presence of 4 mM $Co(NH_3)_6^{3+}$; (x) 0.8 mM A7D hairpin, 10 °C and pH 7; (xi) 1.4 mM A8D hairpin, 15 °C and pH 7−7.5.

3 loop, 5′U**GA**U/3′G**AAG**G, the free energy change is larger for the 7-deazaadenine substitution on the shorter side of the loop, 1.7 kcal/mol, than for the substitutions on the longer side of the loop, 0.6 and 0.9 kcal/mol at 37 °C. When an adenine on the longer side of the loop is changed to a 7-deazaadenine, there exists an alternate adenine with which to form a sheared GA pair. The change in free energy when adenines on the longer side of the loop are modified could be partially explained by a change in conformational flexibility because two possible GA pairs are reduced to one possible GA pair. If the two possible pairings are energetically equivalent, then the entropic effect of this change would make the $\Delta G^\circ_{loop,37}$ less favorable by 0.4 kcal/mol at 37 °C (i.e., RT ln 2). A larger conformational change within the 2 × 3 loop is also possible (see below).

^{1}H NMR Spectroscopy on RNA with Inosine or 7-Deazaadenine Substitutions. Figure 3 shows one-dimensional ^{1}H NMR spectra of two oligonucleotide models of the J4/5 loop from *P. carinii* and several models containing inosine or 7-deazaadenine substitutions. Note that the G8−C13 pair adjacent to one GU closing pair in the duplexes is switched relative to the equivalent G4−C22 pair in the hairpin models. The spectrum of the duplex model of the J4/5 loop from *P. carinii* (Figure 3, i) shows assignments based on one-dimensional NOE difference experiments (*15*). The two most

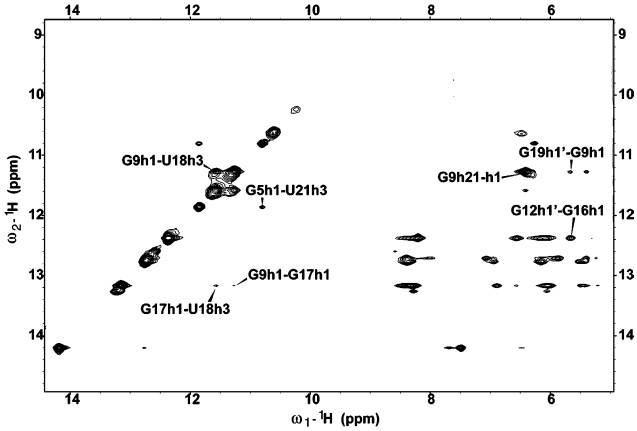

FIGURE 4: 2D SNOESY of the J4/5 hairpin in 80 mM NaCl, 10 mM phosphate, and 0.5 mM Na_2EDTA, pH 5 at 5 °C. The NOESY mixing time is 150 ms.

upfield resonances show no NOEs. The spectrum of the hairpin model (Figure 3, vii) shows imino proton assignments from 2D SNOESY (Figure 4), ^{15}N HSQC, and HNNCOSY experiments (Supporting Information).

With the exception of the A17D substitution in the duplex model, changes in the upfield imino proton resonances (10−12 ppm) suggest changes in the conformations and/or

dynamics of the internal loops with either inosine or 7-deazaadenine substitutions. The imino proton spectra for internal loops with inosine substitutions show two fewer upfield (10−12 ppm) imino proton resonances (Figure 3, ii, iii, viii) and no additional downfield (12−15 ppm) resonance for an inosine imino proton. The absence of the upfield imino proton resonances results from increased solvent proton exchange for imino protons in the closing GU pairs and/or loop nucleotides, which is consistent with the thermodynamic destabilization observed in melting experiments on duplexes with inosine substitutions (Tables 1 and 2). The upfield resonances remain unassigned because no NOEs to these resonances were observed in 1D or 2D NOESY experiments. In a 2D cwfNOESY spectrum (*63*) of the 25-mer hairpin with a G6 → I substitution (G6I hairpin), the G9 and U18 still show a strong cross-peak, indicating formation of the G9−U18 pair, and the G9 imino proton still shows a strong NOE to its own amino proton (data not shown).

The addition of cobalt hexamine to the hairpin construct containing the G6 → I substitution results in a new resonance at about 11 ppm, which could be for the G5−U21 closing pair (Figure 3, ix). Cobalt hexamine can mimic magnesium ions (*64*, *65*). A one-dimensional NOE difference experiment saturating the imino proton resonances showed no NOEs to the hexamine resonance; an NOE difference experiment saturating the hexamine resonance showed NOEs to all but four weak imino protons. This suggests that the addition of cobalt hexamine may stabilize the RNA as a +3 salt rather than as a metal ion binding to a specific site (*64*, *65*). The addition of cobalt hexamine to the unmodified hairpin construct causes aggregation.

The 7-deazaadenine substitutions cause various changes in the imino proton spectra (Figure 3, iv−vi, x, xi). The A6D substitution in the duplex model induces a broadening of the upfield imino proton resonances (Figure 3, vi), which is consistent with its large decrease in thermodynamic stability (Tables 1 and 2). The A16D substitution in the duplex causes changes in the chemical shift and/or intensities of the upfield imino protons. The A17D substitution in the duplex model causes only minor changes in the imino proton spectrum (Figure 3, iv). The same change in the hairpin model of the loop (A8D hairpin), however, results in too many imino proton resonances (Figure 3, xi). The A8D hairpin model also showed 15 rather than 8 H5−H6 cross-peaks in the DQFCOSY spectrum in D_2O, suggesting two stable conformations of the oligonucleotide. The A7D hairpin model showed a shift in the upfield imino proton resonances (Figure 3, x). The chemical shifts for all of the nonexchangeable resonances for the loop nucleotides also change with the A7D substitution in the hairpin. The NOESY, TOCSY, and ROESY spectra for the A7D and unmodified hairpins show cross-peaks indicating chemical exchange on the millisecond time scale or slower. Both hairpins show similar dynamic characteristics.

The 25-mer Hairpin Construct Is in One Global Conformation. The J4/5 hairpin construct was melted in 1 M NaCl, 20 mM sodium cacodylate, and 0.5 mM Na_2EDTA, pH 7, buffer solution at four concentrations ranging from 1.34×10^{-4} to 6.00×10^{-6} M and had a concentration-independent T_m of 76.3 ± 0.5 °C, consistent with formation of a hairpin under these conditions. Gel filtration experiments at 1 M NaCl and 80 mM NaCl show the same elution profile with one peak.

On the basis of predicted and measured diffusion constants for nucleic acids (*66*) and of control measurements on 8-mer Watson−Crick duplexes (*67*), the diffusion constant measured by NMR spectroscopy (*66*), $(0.875 \pm 0.065) \times 10^{-6}$ cm^2/s, is consistent with formation of a 25-mer hairpin; the difference between predicted diffusion constants for a 25-mer hairpin and a 25-mer self-complementary duplex is within experimental error, however. Nevertheless, the diffusion constant value clearly shows that the RNA is not aggregating at the concentration and buffer conditions in which the NMR experiments were performed.

During preparation of the 5′-^{32}P-labeled oligonucleotide for enzymatic mapping experiments, the RNA is purified on a native polyacrylamide gel and shows one band. S1 nuclease mapping shows hits in the tetraloop and the terminal nucleotides, and V1 nuclease mapping shows hits for the Watson−Crick stems, as predicted (*67*). V1 but not S1 nuclease mapping shows hits for the internal loop nucleotides, suggesting that these nucleotides are stacked in the loop.

Proton and Heteronuclear Experiments on the Hairpin Model of the J4/5 Loop from P. carinii. The details of the NMR spectroscopy methods, spectra for several heteronuclear experiments, tables of assignments, restraints for modeling, and the protocol used for modeling are available in Supporting Information. The NOEs in the SNOESY and ^{13}C-edited NOESY provide crucial information about the conformation of the loop and are further discussed below. The results summarized below from several NOESY, ROESY, TOCSY, and ^{13}C HSQC experiments provide evidence for a dynamic loop conformation.

The cross-peaks in the SNOESY spectrum confirm the formation of GU closing pairs and at least one sheared GA pair within the internal loop. The spectrum from the SNOESY experiment at 5 °C and pH 5 shows the sharpest imino and amino proton resonances (Figure 4). The G9 and U18 imino protons show a strong cross-peak typical of GU pairs, although the diagonal peaks for G9 and U18 imino protons are broad and have a distinct shoulder at 15 °C, probably a result of chemical exchange. Both G9 and U18 imino protons show NOEs to the G17 imino proton. The G5 and U21 imino protons show weaker cross-peaks and diagonal peaks than the G9 and U18 imino protons, which suggests that the G5−U21 pair is more dynamic and that these imino protons exchange more with water. Either the U21 or U24 imino proton also shows a weak cross-peak (11.96, 14.20 ppm) due to chemical exchange at 15 °C. The weak diagonal peak at 10.3 ppm is observed only at 5 °C and pH 5 and shows no NOEs. The assignment of this peak to either G6 or G19 is consistent with a guanine imino proton in a sheared GA pair (*68*).

In the H1′ and imino proton (H1) region of the spectrum (5.0−6.5 and 10.0−14.5 ppm), two NOEs, G12 H1′−G16 H1 and G19 H1′−G9 H1, confirm the formation of sheared GA pairs in the GAAA tetraloop and the internal loop, respectively (*46*, *68*). The G9 amino proton (H21) is assigned on the basis of its strong cross-peak to its own imino proton (H1); G9 H21 also shows weak NOEs to C10 H1′ and G19 H1′ (Supporting Information). Thus, the cross-peaks to the

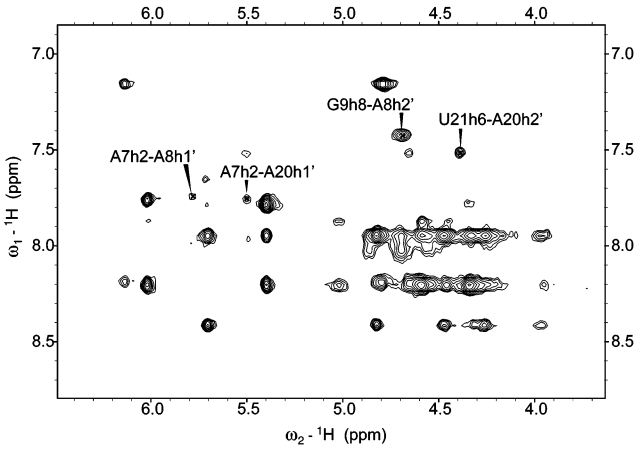

FIGURE 5: 2D ^{13}C-edited NOESY-HSQC at 500 MHz on an adenine-only labeled J4/5 hairpin at 20 °C in 80 mM NaCl, 10 mM phosphate, and 0.5 mM Na_2EDTA, pH 6. The NOESY mixing time is 400 ms. The region shown includes ^{13}C-labeled ribose to aromatic resonances.

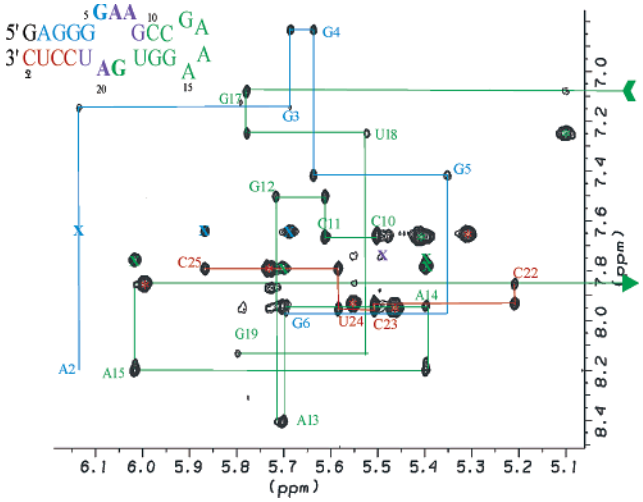

FIGURE 6: 2D NOESY H1′−H6/H8 region of the J4/5 hairpin at 20 °C, 400 ms mixing time, and 750 MHz in 80 mM NaCl, 10 mM phosphate, and 0.5 mM Na_2EDTA, pH 6. The intranucleotide H1′−H8 cross-peak is labeled. The purine walk is in blue, the walk through the tetraloop to G19 is in green, and the pyrimidine walk is in red. Asterisks mark the H5−H6 cross-peaks; X's mark the H2−H1′ cross-peaks.

G9 imino and amino protons significantly restrain the nucleotides in the GU and GA pairs in this region of the loop.

The cross-peaks in a ^{13}C-edited NOESY−HSQC spectrum of a sample with only adenines isotopically labeled provide important NOEs for the internal loop and confirm that A8 and A20 probably stack on the adjacent helices (Figure 5). Although U21 H6 and G9 H8 resonances are overlapped with other aromatic resonances, strong cross-peaks from A8 H2′ to G9 H8 and A20 H2′ to U21 H6 clarify assignments of the internal loop nucleotides. No G19 H1′ or H2′ to A20 H8 cross-peak is observed, although a weak G19 H3′ to A20 H8 cross-peak is observed at a mixing time of 400 ms. The cross-peaks in the ribose region of the HCCHTOCSY on the adenine -labeled sample (Supporting Information) confirmed the assignment of sugar protons in A8 and A20.

Assignment of the internal loop adenine H2 and H8 resonances is based on several weak NOEs in the 400 ms mixing time ^{13}C-edited NOESY−HSQC (Figure 5) and the following assumptions based on NMR spectroscopy and biochemical results: (1) the assigned weak NOEs in the H1′−H8/H2 region originate from one conformation of the loop and are not the result of chemical exchange or spin diffusion; (2) all of the loop nucleotides are stacked because the loop nucleotides are recognized by V1 and not S1 nuclease (*67*); (3) sheared GA pairs are likely because the inosine and 7-deazaadenine substitutions destabilize the loop thermodynamically and structurally; (4) cross-strand NOEs to adenine H2 resonances are possible if sheared purine pairs are present.

The assignments for A7 and A8 H2, H8, and H1′ protons, however, lack unambiguous confirmation. The A7 and A8 resonances are broad. Chemical exchange for the A7 and A8 H2 resonances complicates interpretation of weak NOEs; a weak peak could result from two protons very near each other in a minor conformation. Spin diffusion is another possible source of weak NOEs in experiments with 400 ms mixing times. Adenine HCCHTOCSY and adenine HNCTOCSYCH experiments can confirm assignments of adenine H2 and H8 protons by through-bond correlations but did not provide any information on the adenines in the internal loop due to relaxation effects [data not shown (*67*)]. Thus, many of the NOEs used for assignments were not used in the final modeling. Nonetheless, the current set of assignments is the best possible interpretation of the spectrum that is consistent with all of the available biochemical and NMR data.

The crystal structure of the J4/5 loop from *T. thermophila* (*8*) provides insight into the NOEs that are possible if the J4/5 loop in *P. carinii* forms a similar structure (Supporting Information). The H2 resonance at 7.75 ppm was assigned to A7 H2 because A7 H2 is the proton most likely to give NOEs to both A20 H1′ and A8 H1′. The H2 resonance at 7.61 ppm was assigned to A20 H2 because an H2−H2 NOE (7.61, 7.75 ppm) observed at 100 ms is consistent with a cross-strand stack that is common in adjacent sheared purine pairs; thus, A20 H2 is the proton most likely to give a medium NOE to A7 H2 if a cross-strand stack exists (*8*) (Supporting Information). In contrast, an NOE between consecutive stacked adenines, i.e., A7 H2−A8 H2, is expected to be weak, as previously observed in GAAA tetraloops (*69*, *70*).

The H1′−H8/H6 NOESY walk is shown in Figure 6. The NOESY walk is discontinuous in the internal loop after G6 and G19. In tandem sheared GA pairs, the H1′−H8/H6 walk is also discontinuous after the guanine (*68*). Some weak cross-peaks in the internal loop, such as U18 H1′−G19 H8, are observed at 20 °C, but not at higher temperatures, suggesting that the loop structure becomes more dynamic at higher temperatures. As the temperature is increased, more cross-peaks due to chemical exchange and more weak H1′−H2′ cross-peaks are observed in the ROESY and TOCSY spectra, respectively (Supporting Information), also suggesting that the loop becomes more dynamic at higher temperatures. Although there is evidence for chemical exchange on the millisecond time scale or slower for H1′, H5, H6, and H2 resonances in the internal loop and terminal nucleotides (see Supporting Information), the relative intensities of these pairs of peaks suggest that one conformation predominates. If two or more equally populated conforma-

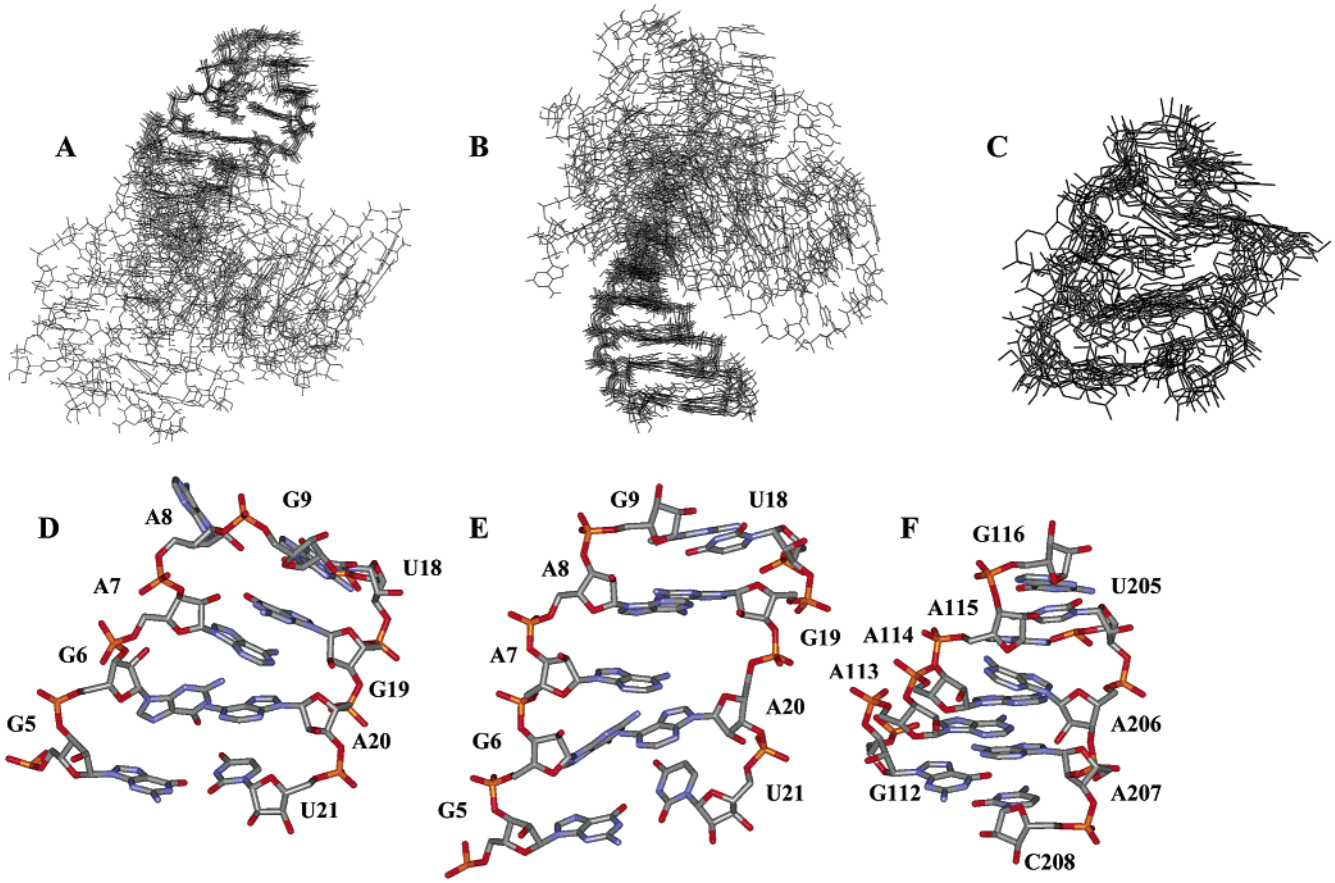

FIGURE 7: Visual aids for discussion of possible structures: (A) 12 molecules with the upper stem and hairpin nucleotides superimposed. (B) The same 12 molecules with the lower stem nucleotides superimposed. (C) Superposition of 12 possible internal loop structures. (D) One of 15 molecules with G19−A7 hydrogen bond restraints that satisfies 9 of 12 weak internucleotide NOEs not included as restraints. (E) One of 15 molecules with G19−A8 hydrogen bond restraints that satisfies 9 of 12 weak internucleotide NOEs not included as restraints, including A7 H2−A20 H1′ and A7 H2−A8 H1′. (F) Nucleotides 112−116 and 206−210 from the crystal structure of the P4−P6 region of the *T. thermophila* group I intron (1GID.pdb in the Protein Data Bank) (*8*).

tions existed in solution, then the intensities of the resonances in slow chemical exchange would be equal at both chemical shifts.

Modeling Shows That More Than One Conformation Satisfies the Experimental Restraints. A set of 15 molecules with G19−A8 and G6−A20 hydrogen bond restraints, 15 molecules with G19−A7 and G6−A20 hydrogen bond restraints, and 3 molecules with no hydrogen bond restraints for sheared GA pairs show a reasonable global fold with distance violations not greater than 0.2 Å and with low total energies. Figure 7A,B shows the overlay of 12 low free energy structures with either the lower stem or the upper stem and hairpin loop superimposed. Table 3 lists for each nucleotide the RMSD in the superimposed structures and the number of distance restraints used in the modeling. The nucleotides in the stem and hairpin loop have a sufficient number of restraints to be well-defined in the models.

The 62 NOEs for the 9 internal loop nucleotides are insufficient to define a unique conformation of the internal loop, to determine unambiguously the base pairing in the loop, or to restrain the orientation of the helices during the molecular dynamics simulations. The lack of NOEs, especially cross-strand NOEs, for the loop nucleotides allows many conformations in the internal loop (Figure 7C). G6, G19, and A20 probably stack on the helix ends, but the interface between the two helices is poorly defined.

Panels D and E of Figure 7 are visual aids for the discussion of possible structures of the loop. One prediction for the loop structure is that the J4/5 loop in *P. carinii* is similar to the J4/5 loop in *T. thermophila* and has sheared GA base pairs between G6−A20 and G19−A7. Another possible prediction for loop structure is that base pairs form between G6−A20 and G19−A8; this base pairing arrangement would be consistent with the models used to predict the thermodynamic stabilities of 2 × 3 internal loops and the unusual stability of the J4/5 loop from *P. carinii* (*15*, *79*). These predictions were tested by modeling with hydrogen bond restraints for each possible base pairing. Panels D and E of Figure 7 show two possible loop conformations, one example for each set of 15 molecules with G6−A20 and either G19−A7 (Figure 7D) or G19−A8 (Figure 7E) hydrogen bond restraints. Both possible base-pairing arrangements are consistent with the experimental NOE restraints. In both loops, the N3 atoms of A7 and A20 are pointed into the minor groove and could act as hydrogen bond acceptors. The positions of the adenine N3 functional groups in these two possible conformations of the *P. carinii* loop are similar to those in the J4/5 loop in *T. thermophila* (Figure 7F) (*7*−*10*).

DISCUSSION

The J4/5 loop in the mouse-derived *P. carinii* group I intron has an important biological role in rRNA processing

Table 3: Averaged RMSD Values and the Number of Intra- and Internucleotide Restraints for Each Nucleotide in the J4/5 Hairpin[a]

	averaged RMSD		total restraints	distance restraints	
nucleotide	lower stem[b]	upper stem, hairpin loop[c]	per nucleotide	intra-nucleotide	inter-nucleotide
G1	3.60		3	0	3
A2	1.71		8	1	7
G3	1.50		18	4	14
G4	1.34		22	7	15
G5	1.56		19	7	12
G6	4.38	10.66	6	3	3
A7	10.72	6.15	2	0	2
A8	16.08	5.53	5	3	2
G9		1.57	11	0	11
C10		0.98	20	5	15
C11		1.01	26	6	20
G12		1.21	18	6	12
A13		1.65	20	11	9
A14		1.35	30	12	18
A15		1.27	25	10	15
G16		0.75	25	6	19
G17		0.90	24	6	18
U18		1.28	20	6	14
G19	11.13	1.96	10	4	6
A20	4.92	4.69	6	3	3
U21	1.87		8	2	6
C22	1.76		17	7	10
C23	1.24		18	4	14
U24	1.27		20	6	14
C25	2.90		13	5	8

[a] RMSD values were determined by averaging the pairwise RMSD values for each nucleotide. [b] The averaged RMSD values of the lower stem nucleotides with the stem G1-G6 and A20-C25 nucleotides superimposed. [c] The averaged RMSD values of the upper stem and tetraloop nucleotides with the G9-A20 nucleotides superimposed.

in a fungal pathogen and thus models a potential target for rational drug design. This internal loop is also more thermodynamically stable than other 2 × 3 loops with two GU closing pairs (*15*). NMR spectra indicate that the internal loop is conformationally dynamic but predominantly in one average conformation; i.e., the loop gives rise to one predominant set of chemical shifts that represent one structure or an ensemble average of structures. The thermodynamic stabilities, NMR data, and the effects of inosine and 7-deazaadenine substitutions all contribute to a model of the possible conformations of the J4/5 loop in *P. carinii*.

The Effects of Inosine and 7-Deazaadenine Substitutions on the Thermodynamic Stabilities and Structures of the Loop Are Consistent with Formation of Sheared GA Pairs. As shown in Tables 1 and 2, substitution of inosine for a guanine or 7-deazaadenine for an adenine in the internal loop is thermodynamically destabilizing. Similar effects have been observed for substitutions of sheared GA pairs in GCAA and GAAA tetraloops (*59*, *71*) and hammerhead ribozymes (*58*, *72*, *73*). If the nucleotides in the loop formed imino GA pairs, then inosine or 7-deazaadenine substitutions are predicted to increase the thermodynamic stability because the pK_a change favors hydrogen bonding. For G → I substitutions, fewer imino proton resonances are observed for the loop nucleotides, indicating increased solvent exposure suggesting that the loop becomes less structured. With the exception of the A17D duplex, A → D substitutions change the imino proton spectra of the J4/5 loop. When A7 in the hairpin model is a 7-deazaadenine, the chemical shifts of the nonexchangeable protons and the pattern of NOEs to the internal loop nucleotides change. Thus, when the functional groups that form hydrogen bonds in sheared GA pairs are changed, the structure and thermodynamic stability of the loop change.

NMR Spectra Provide Evidence for the Formation of at Least One Sheared GA Pair in the Internal Loop. The upfield chemical shift of the weak imino proton resonance assigned to either G6 or G19 (near 10.3 ppm, Figure 4) in the hairpin model is consistent with the imino proton chemical shifts in sheared GA pairs (*68*, *71*); in contrast, imino protons in imino GA pairs resonate near 11.5−12.5 ppm. A strong guanine H1 to adenine H2 NOE is expected in an imino GA pair but is not observed. The NOE from G19 H1′ to G9 H1 is also consistent with the formation of a sheared GA pair but not an imino GA pair (*46*, *68*, *74*).

The NMR Data and Structural Modeling Are Consistent with a Dynamic Internal Loop. The few NOE distance restraints observed for the internal loop can be satisfied with structures generated using hydrogen bond restraints for several different combinations of sheared purine pairs. This even includes the possibility of a sheared AA and GA pair, although this is not consistent with the effects of inosine substitutions. When more than one conformation exists, NOE distances do not contain sufficient information to define a unique structure (*75*). Even with the added hydrogen-bonding restraints for one set of GA pairs, the entire molecule does not converge to one structure, although superposition of the lower and upper helices separately does show convergence. The loop nucleotides probably stack on the stem helix ends, but the interface between the helices is poorly defined. The lack of cross-strand NOEs and long-range restraints (*47*, *76*) is evident in the many different orientations of the upper and lower helices that result from the molecular dynamics simulations.

The J4/5 loop in *P. carinii* probably does not have a single, static structure. At 20 °C, the ROESY spectra show positive evidence for chemical exchange for A7, A8, and A20 H2 resonances on the millisecond time scale or slower (Supporting Information). Many of the resonances for the nucleotides in the internal loop are broad and weak. The very few and very weak NOEs to loop nucleotides, especially A7, are consistent with a dynamic internal loop. At temperatures higher than 20 °C, the NMR data show fewer NOEs to loop nucleotides in NOESY spectra, more cross-peaks due to chemical exchange in ROESY and TOCSY spectra, and more H1′−H2′ cross-peaks in TOCSY spectra indicating conformational dynamics in the ribose sugar puckers for loop nucleotides, while the evidence for the duplex stems shows less change. Thus the loop nucleotides show even more evidence of conformational dynamics at higher temperatures. Nonetheless, the relative intensities of the pairs of resonances showing evidence for slow chemical exchange suggest that one average conformation predominates at 20 °C. The lack of NOE restraints for the loop nucleotides, however, prohibits establishing a well-defined model of the predominant average conformation.

The loop is partly structured, however. A model of a completely unstructured loop is inconsistent with the following data: (1) NOEs in the NOESY walk suggesting that G6, G19, and A20 stack on the adjacent stem helices, (2) changing one functional group on an adenine or guanine changes the stability and structure of the loop, and (3) the

loop nucleotides are recognized by V1 and not S1 nuclease (*67*). More than one stable conformation of loop may occur; the loop may be both partly structured and dynamic. When no hydrogen bond restraints are used for the loop nucleotides, several possible structures still show stacked bases and possible pairing.

One Possible Conformation of the J4/5 Loop from P. carinii Is Consistent with the Base Pairing Observed in the J4/5 Loop from T. thermophila. Figure 7D shows one possible conformation of the J4/5 loop from *P. carinii* with sheared GA pairs, G6−A20 and G19−A7. The N3 atoms of A20 and A7 (A207 and A114 in *T. thermophila*, Figure 7F) are pointed into the minor groove and are available as hydrogen bond acceptors. The position of A20 2′OH in the minor groove is compatible with all of the experimental NOE restraints. The N3 atoms of A207 and A114 and the 2′OH of A207 have been identified in NAIM and NAIS analysis as being important for the docking interaction (*3*, *9*, *28*). Thus, one possible conformation of the loop supports the hypothesis that internal loops with the same function conserve the same shape (*77*, *78*).

The J4/5 Loop from P. carinii Is Relatively Stable Thermodynamically Even Though It May Adopt More Than One Base Pairing Conformation. At 37 °C, the J4/5 loop from *P. carinii* is 2.2 kcal/mol more favorable than the equivalent loop with all A nucleotides, 5′ U**AA**U/3′G**AAA**G (*15*). The rules for predicting thermodynamic stabilities of 2 × 3 loops attribute this additional stability to formation of GA pairs adjacent to the helix ends, i.e., G6−A20 and G19−A8 pairs (*79*). Figure 7E shows one possible conformation with this base pairing. G19 may instead form a sheared GA pair with A7. Both possible sets of GA pairs present hydrogen bond acceptors in the minor groove in a similar way as observed in the *T. thermophila* J4/5 loop (Figure 7). Exchange between several stable conformations of the loop may contribute to its thermodynamic stability. The cross-strand NOEs to G19 define its stacked position on the G9−U18 closing pair. The pattern of NOEs suggests that G6, G19, A8, and A20 may stack on the helix stems. Thus both NMR and chemical substitution data are consistent with sheared GA pairs enhancing the thermodynamic stability of the *P. carinii* J4/5 loops, consistent with the rules for predicting thermodynamic stabilities of 2 × 3 loops.

The J4/5 Loop from P. carinii Shows Different Characteristics Than Other Loops with Similar Sequences. Results for the J4/5 loop can be compared with those for similar loops that are also involved in tertiary interactions. An NMR analysis of an internal loop, 637C**GA**G640/623G**AAG**C619, at the Varkud satellite RNA active site provides unambiguous experimental evidence for the formation of hydrogen bonds in two sheared GA pairs (G620−A639 and G638−A621) in an NMR structure very well defined by NOE restraints (*80*). G638, A621, and A622 form a shared sheared GA pair structure. This structure suggests another possible conformation of the J4/5 loop in *P. carinii*. A shared sheared GA pair conformation is consistent with the effects of inosine and 7-deazaadenine substitutions and with the NOEs observed in the J4/5 internal loop.

Interestingly, the 5′C**GA**G/3′G**AAG**C loop shows no NMR evidence for dynamics on the millisecond time scale or slower. This contrasts with the 5′U**GA**U/3′G**AAG**G loop. Replacing GU with more stable GC closing pairs is one possible cause for the change in dynamics. The other differences in stem sequences may also affect dynamics. The J4/5 hairpin studied here has a lower stem of sequential guanines paired with sequential pyrimidines. Alternate pairings with G5 and G6 are possible in the model for the J4/5 loop but not the 5′C**GA**G/3′G**AAG**C loop. In addition to being more structurally rigid, the 5′C**GA**G/3′G**AAG**C loop is also more thermodynamically stable than the 5′U**GA**U/3′G**AAG**G loop (Table 2).

In the crystal structure of the *Thermus thermophilus* 30S ribosomal subunit (*81*), an internal loop with the same sequence is found in the 16S rRNA helix 44, 1466C**GA**G1469/1435G**AAG**C1431. This loop forms part of an intersubunit bridge in the 70S ribosome. The crystal structure shows three adenines stacking and G1467 and G1432 in approximately the correct orientation to form sheared GA pairs. G1467 and G1432, however, are not coplanar with the respective adenines, and except for G1467 N2−A1433 N7, the distances between heavy atoms are longer than 3.2 Å and thus characteristic of only weak, electrostatic hydrogen bonds. The NMR and X-ray crystal structures of this loop are very similar (*80*) and provide further evidence for the structural stability of this loop motif.

A similar DNA internal loop, 5′C**GA**G/3′G**AAG**C, is found at the 3′ terminus of the parvovirus genome. Its NMR structure shows an unusual cross-strand stack of three purines (*82*), which has not been observed in RNA loops. This DNA internal loop also has some γ angles close to trans, which give rise to the downfield-shifted phosphorus resonances observed (*82*). This loop is more thermodynamically stable than other DNA loops with T, C, or G as the middle nucleotide.

Asymmetric internal loops in models of the HIV-1 package signal contain imino GA pairs. The structural and thermodynamic effects of inosine substitutions in those cases, however, differ from the effects observed in models of the J4/5 loop (*83*). The NMR evidence and rationale for imino GA pairs seen in other asymmetric internal loops (*83*−*85*) are not observed in the *P. carinii* J4/5 loop.

Interestingly, there is no evidence that an imino GA pair stabilizes the J4/5 loop. The 2 × 2 loop, 5′G**GA**U/3′U**AG**G, has two imino GA pairs but is 1.7 kcal/mol less stable than a 2 × 2 loop, 5′U**GA**G/3′G**AG**U, with two sheared GA pairs (*86*). The loop 5′G**AG**U/3′U**GA**G also has two imino GA pairs and is even less stable (*15*). The least stable symmetric 2 × 2 loop with GA pairs and GU closing pairs, 5′U**AG**G/3′G**GA**U, has no observable imino proton resonance indicating the conformation of the GA pair (*15*). Evidently, imino GA pairs adjacent to GU pairs are relatively unstable.

The Conformational Dynamics of an Asymmetric Internal Loop May Be an Important Factor in the Biological Function of the Loop and May Facilitate Drug Discovery. The conformational dynamics of the J4/5 loop may be an important aspect of its function; the loop may adopt one conformation to bind the P1 helix during the first step of the self-splicing reaction and then change its conformation to reach the transition state or facilitate the second step of splicing. A roughly isoenergetic conformational shift to the transition state would be beneficial because catalysis is facilitated by formation of new interactions in the transition state without a net loss of interactions in the ground state. Moreover, specificity in formation of tertiary interactions

may be enhanced because the internal loop is flexible. Binding the P1 helix may have an unfavorable entropy change associated with forcing the J4/5 loop into one conformation. Overcoming this unfavorable free energy change may require several favorable interactions, thus providing specificity. Binding to a completely unstructured loop, however, would be associated with a larger unfavorable entropy change requiring more favorable interactions. Thus a partially structured loop may represent a compromise that allows specificity without requiring extensive complexity from many interactions.

Both GU and sheared GA pairs present a unique configuration of functional groups with the potential to form hydrogen bonds. A specific pattern of hydrogen bond donors and acceptors in a loop is a good target for rationally designing a drug with high specificity. The dynamics of an asymmetric internal loop may also be beneficial for designing and screening drugs. A structurally dynamic loop may present more possible options for binding. Thus, loops similar to the J4/5 loop from the *P. carinii* group I intron may be attractive drug targets.

SUMMARY

The J4/5 internal loop from the *P. carinii* group I intron is a thermodynamically stable and dynamic loop with more than one possible combination of sheared GA pairs. The formation of sheared GA pairs is consistent with NMR spectra and the effects of inosine and 7-deazaadenine substitutions. One possible conformation of the loop is similar to the sheared adenine pairs in the J4/5 loop in *T. thermophila*; however, alternative base pairings also present critical functional groups in a similar orientation. The formation of sheared GA pairs adjacent to GU closing pairs and the stacking of the loop nucleotides on the adjacent helices contribute to the thermodynamic stability of the loop.

ACKNOWLEDGMENT

The authors thank Dr. Thomas Barnes and Dr. Tianbing Xia for assistance with the preparation of isotopically labeled NTPs as well as several laboratories in the Biology and Chemistry Departments at the University of Rochester for the use of their equipment in this protocol. The authors thank Prof. Thomas Shrader for the plasmid for the His-tag T7 enzyme. The authors also thank the following people for advice and assistance on the protocol for preparing T7: Prof. Kevin Weeks, Dr. Ivelitza Garcia, Dr. David Mathews, Joshua Diamond, and Jeffrey Sabina. The authors thank Prof. Ignacio Tinoco, Jr., Dr. Minxue Zheng, Dr. Chul-Hyun Kim, and Dr. Rueben Gonzalez at the University of California, Berkeley, for NMR time, assistance in data collection, and discussions. The authors also thank Prof. Tinoco and Dr. Kevin Luebke for sharing unpublished spectra of the J4/5 loop from *T. thermophila*. The authors thank Dr. Jack Skalicky for assistance with data collection at the 750 MHz NMR facility at the University of Buffalo.

SUPPORTING INFORMATION AVAILABLE

A table of parameters for NMR spectroscopy experiments; tables of proton, carbon, nitrogen, and phosphorus chemical shifts; a table of distance restraints; a table of distances between protons added to the crystal structure of the J4/5 loop in *T. thermophila*; 10 figures of NMR spectra; and a description of the simulated annealing protocol. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Michel, F., and Westhof, E. (1990) *J. Mol. Biol. 216*, 585−610.
2. Strobel, S. A., and Cech, T. R. (1995) *Science 267*, 675−679.
3. Strobel, S. A., and Ortoleva-Donnely, L. (1999) *Chem. Biol. 5*, 153−165.
4. Bevilacqua, P. C., Kierzek, R., Johnson, K. A., and Turner, D. H. (1992) *Science 258*, 1355−1358.
5. Herschlag, D. (1992) *Biochemistry 31*, 1386−1399.
6. Pyle, A. M., Murphy, F. L., and Cech, T. R. (1992) *Nature 358*, 123−128.
7. Cate, J. H., Gooding, A. R., Podell, E., Zhou, K., Golden, B. L., Szewczak, A. A., Kundrot, C. E., Cech, T. R., and Doudna, J. A. (1996) *Science 273*, 1696−1699.
8. Cate, J. H., Gooding, A. R., Podell, E., Zhou, K., Golden, B. L., Szewczak, A. A., Kundrot, C. E., Cech, T. R., and Doudna, J. A. (1996) *Science 273*, 1678−1685.
9. Strobel, S. A., Ortoleva-Donnely, L., Ryder, S., Cate, J. H., and Moncoeur, E. (1998) *Nat. Struct. Biol. 5*, 60−66.
10. Ortoleva-Donnely, L., Szewczak, A. A., Gutell, R. R., and Strobel, S. A. (1998) *RNA 4*, 498−519.
11. Cech, T. R., Tanner, K., Tinoco, I., Jr., Weir, B., Zuker, M., and Perlman, P. (1983) *Proc. Natl. Acad. Sci. U.S.A. 80*, 3903−3907.
12. Inoue, T., and Cech, T. R. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 648−652.
13. Zaug, A., and Cech, T. R. (1995) *RNA 1*, 363−374.
14. Lehnert, V., Jaeger, L., Michel, F., and Westhof, E. (1996) *Chem. Biol. 3*, 993−1009.
15. Schroeder, S. J., and Turner, D. H. (2001) *Biochemistry 40*, 11509−11517.
16. Disney, M. D., Haidaris, C. G., and Turner, D. H. (2001) *Biochemistry 40*, 6507−6519.
17. Disney, M. D., Gryaznov, S. M., and Turner, D. H. (2000) *Biochemistry 39*, 14269−14278.
18. Testa, S. M., Haidaris, C. G., Gigliotti, F., and Turner, D. H. (1997) *Biochemistry 36*, 15303−15314.
19. Testa, S. M., Gryaznov, S. M., and Turner, D. H. (1998) *Biochemistry 37*, 9379−9385.
20. Testa, S. M., Gryaznov, S. M., and Turner, D. H. (1999) *Proc. Natl. Acad. Sci. U.S.A. 96*, 2734−2739.
21. Testa, S. M., Disney, M. D., Turner, D. H., and Kierzek, R. (1999) *Biochemistry 38*, 16655−16662.
22. Disney, M. D., Testa, S. M., and Turner, D. H. (2000) *Biochemistry 39*, 6991−7000.
23. Disney, M. D., Matray, T., Gryaznov, S. M., and Turner, D. H. (2001) *Biochemistry 40*, 6520−6526.
24. Childs, J. L., Disney, M. D., and Turner, D. H. (2002) *Proc. Natl. Acad. Sci. U.S.A. 99*, 11091−11096.
25. Mei, H., Lemrow, S., and Czarnik, A. (1997) *Bioorg. Med. Chem. 5*, 1185−1195.
26. Miletti, K., and Leibowitz, M. (2000) *Antimicrob. Agents Chemother. 44*, 958−966.
27. Damberger, S. H., and Gutell, R. R. (1994) *Nucleic Acids Res. 22*, 3508−3510.
28. Soukup, J., Minakawa, N., Matsuda, A., and Strobel, S. A. (2002) *Biochemistry 41*, 10426−10438.
29. Ortoleva-Donnely, L., Kronman, M., and Strobel, S. A. (1998) *Biochemistry 37*, 12933−12942.
30. Strobel, S. A., and Shetty, K. (1997) *Proc. Natl. Acad. Sci. U.S.A. 94*, 2903−2908.
31. Scaringe, S. A., Franklyn, C., and Usman, N. (1990) *Nucleic Acids Res. 18*, 5433−5441.
32. Usman, N., Ogilvie, K. K., Jiang, M. Y., and Cedergren, R. J. (1987) *J. Am. Chem. Soc. 109*, 7845−7854.
33. Puglisi, J., and Wyatt, J. (1995) *Methods Enzymol. 261*, 323−350.
34. Milligan, J., Groebe, D., Witherell, G. W., and Uhlenbeck, O. C. (1987) *Nucleic Acids Res. 21*, 8783−8798.
35. Kao, C., Zheng, M., and Rudisser, S. (1999) *RNA 5*, 1268−1272.
36. Batey, R., Inada, M., Kujawinski, E., Puglisi, J., and Williamson, J. (1992) *Nucleic Acids Res. 20*, 4515−4523.
37. Batey, R., Battiste, J., and Williamson, J. (1995) *Methods Enzymol. 261*, 300−322.

38. Kim, J. (1997) Ph.D. Thesis, Department of Chemistry, University of Rochester, Rochester, NY.
39. Borer, P. N. (1975) in *Handbook of Biochemistry and Molecular Biology: Nucleic Acids* (Fasman, G. D., Ed.) p 589, CRC Press, Cleveland, OH.
40. Petersheim, M., and Turner, D. H. (1983) *Biochemistry 22*, 256−263.
41. Borer, P. N., Dengler, B., Tinoco, I., Jr., and Uhlenbeck, O. C. (1974) *J. Mol. Biol. 86*, 843−853.
42. McDowell, J. A., and Turner, D. H. (1996) *Biochemistry 35*, 14077−14089.
43. McDowell, J. A. (1996) Ph.D. Thesis, Department of Chemistry, University of Rochester, Rochester, NY.
44. Smallcombe, S. (1993) *J. Am. Chem. Soc. 115*, 4776−4785.
45. Hore, P. J. (1983) *J. Magn. Reson. 55*, 283−300.
46. Lukavsky, P., and Puglisi, J. D. (2001) *Methods 25*, 316−332.
47. Hennig, M., Williamson, J. R., Brodsky, A., and Battiste, J. (2000) *Curr. Protocols Nucleic Acid Chem. 7*, 1−30.
48. Varani, G., and Tinoco, I., Jr. (1991) *Q. Rev. Biophys. 24*, 479−532.
49. Varani, G., Aboul-ela, F., and Allain, F. H.-T. (1996) *Prog. Nucl. Magn. Reson. Spectrosc. 29*, 51−127.
50. Wijmenga, S., Mooren, M., and Hilbers, C. (1993) in *NMR of Macromolecules: A Practical Approach* (Roberts, G., Ed.) pp 217−288, Oxford University Press, Oxford.
51. Anderson (1983) *J. Comput. Phys. 52*, 24−34.
52. Cornell, W., Cieplak, P., Bayly, C., Gould, I., Merz, K., Ferguson, D., Spellmeyer, D., Fox, T., Caldwell, J., and Kollman, P. (1995) *J. Am. Chem. Soc. 117*, 5179−5197.
53. Freier, S. M., Sugimoto, N., Sinclair, A., Alkema, D., Neilson, T., Kierzek, R., Caruthers, M. H., and Turner, D. H. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 9373−9377.
54. Freier, S. M., Sugimoto, N., Sinclair, A., Alkema, D., Neilson, T., Kierzek, R., Caruthers, M. H., and Turner, D. H. (1986) *Biochemistry 25*, 3214−3219.
55. Turner, D. H., Sugimoto, N., and Freier, S. M. (1988) *Annu. Rev. Biophys. Biophys. Chem. 17*, 167−192.
56. Mathews, D. H., Sabina, J., Zuker, M., and Turner, D. H. (1999) *J. Mol. Biol. 288*, 911−940.
57. SantaLucia, J., Jr., Kierzek, R., and Turner, D. H. (1991) *J. Am. Chem. Soc. 113*, 4313−4322.
58. Seela, F., Mersmann, Grabsky, J., and Gait, M. (1993) *Helv. Chim. Acta 76*, 1809−1820.
59. SantaLucia, J., Jr., Kierzek, R., and Turner, D. H. (1992) *Science 256*, 217−219.
60. Turner, D. H., Sugimoto, N., Kierzek, R., and Dreiker, S. D. (1987) *J. Am. Chem. Soc. 109*, 3783−7385.
61. SantaLucia, J. J. (1991) Ph.D. Thesis, Department of Chemistry, University of Rochester, Rochester, NY.
62. Carter, B., SantaLucia, J., Jr., Turner, D. H., and Holbrook, S. (1997) *Nucleic Acids Res. 25*, 4117−4122.
63. Chen, X., Kierzek, R., and Turner, D. H. (2001) *J. Am. Chem. Soc. 123*, 1267−1274.
64. Kieft, J., and Tinoco, I., Jr. (1997) *Structure 5*, 713−721.
65. Gonzalez, R., and Tinoco, I., Jr. (2001) *Methods Enzymol. 338*, 421−443.
66. Lapham, J., Rife, J., Moore, P. B., and Crothers, D. M. (1997) *J. Biomol. NMR 10*, 255−262.
67. Schroeder, S. J. (2002) Ph.D. Thesis, Department of Chemistry, University of Rochester, Rochester, NY.
68. SantaLucia, J., Jr., and Turner, D. H. (1993) *Biochemistry 32*, 12612−12623.
69. Jucker, F., Heus, H., Yip, P., Moors, E., and Pardi, A. (1996) *J. Mol. Biol. 264*, 968−980.
70. Rudisser, S., and Tinoco, I., Jr. (2000) *J. Mol. Biol. 295*, 1211−1223.
71. Heus, H., and Pardi, A. (1991) *Science 253*, 191−194.
72. Seela, F., Debelak, H., Usman, N., Burgin, A., and Beigelman, L. (1998) *Nucleic Acids Res. 26*, 1010−1018.
73. Fu, D.-J., and McLaughlin, L. (1992) *Biochemistry 31*, 10941−10949.
74. Wu, M., and Turner, D. H. (1996) *Biochemistry 35*, 9677−9689.
75. Bonvin, A., Brunger, A. (1996) *J. Biomol. NMR 7*, 72−76.
76. Hansen, M., Mueller, L., and Pardi, A. (1998) *Nat. Struct. Biol. 5*, 1065−1074.
77. Leontis, N., Stombaugh, J., and Westhof, E. (2002) *Nucleic Acids Res. 30*, 3497−3531.
78. Gautheret, D., Konings, D., and Gutell, R. R. (1994) *J. Mol. Biol. 242*, 1−8.
79. Schroeder, S. J., and Turner, D. H. (2000) *Biochemistry 39*, 9257−9274.
80. Hoffmann, B., Mitchell, G. T., Gendron, P., Major, F., Andersen, A. A., Collins, R. A., and Legault, P. (2003) *Proc. Natl. Acad. Sci. U.S.A. 100*, 7003−7008.
81. Wimberly, B. T., Broderson, D. E., Clemons, W. M., Jr., Morgan-Warren, R. J., Carter, A. P., Vonrhein, C., Hartsch, T., and Ramakrishnan, V. (2000) *Nature 407*, 327−339.
82. Chou, S.-H., and Chin, K.-H. (2001) *J. Biomol. NMR 21*, 307−319.
83. Greatorex, J., Gallego, J., Varani, G., and Lever, A. (2002) *J. Mol. Biol. 322*, 543−557.
84. Lawrence, D. C., Stover, C. C., Noznitsky, J., Wu, Z., and Summers, M. F. (2003) *J. Mol. Biol. 326*, 529−542.
85. Yuan, Y. Q., Kerwood, D. J., Paoletti, A. C., Shubsda, M. F., and Borer, P. N. (2003) *Biochemistry 42*, 5259−5269.
86. Walter, A. E., Wu, M., and Turner, D. H. (1994) *Biochemistry 33*, 11349−11354.
87. Gralla, J., and Crothers, D. M. (1973) *J. Mol. Biol. 78*, 301−319.
88. Xia, T., SantaLucia, J. J., Jr., Burkard, M. E., Kierzek, R., Schroeder, S. J., Jiao, X., Cox, C., and Turner, D. H. (1998) *Biochemistry 37*, 14719−14735.
89. Bloomfield, V., Crothers, D., and Tinoco, I., Jr. (2000) *Nucleic Acids: Structures, Properties, and Functions*, University Science Books, Sausalito, CA.
90. Burkard, M. E., Turner, D. H., and Tinoco, I., Jr. (1998) *The RNA World* (Gesteland, R. F., Cech, T. R., and Atkins, J. F., Eds.) pp 675−680, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

BI0301587