An NMR Study of the HIV-1 TAR Element Hairpin†

John A. Jaeger‡ and Ignacio Tinoco, Jr.*

Department of Chemistry and Laboratory of Chemical Biodynamics, University of California, Berkeley, California 94720

Received June 30, 1993; Revised Manuscript Received August 30, 1993®

ABSTRACT: The TAR hairpin is an important part of the 5′ long terminal repeat of HIV-1 and appears to be recognized by a cellular protein. A 14-base model of the native TAR hairpin 5′-GAGC[CUGGGA]-GCUC-3′ (loop bases in square brackets) has been studied by proton, phosphorus, and natural abundance carbon NMR; these results are compared to other published NMR studies of the TAR hairpin. Assignments of all nonexchangeable protons and of all the stem-exchangeable protons have been made, as well as all phosphorus and many carbon resonances. Large $J_{1'2'}$ and $J_{3'4'}$ proton–proton coupling in the C5, G8, and G9 sugars indicate an equilibrium between C2′- and C3′-*endo* forms; these data show a dynamic loop structure. We see three broad imino resonances that have not been reported before; these resonances are in the right region for unbonded loop imino protons. These peaks suggest the protons are protected from fast exchange with the solvent by the structure of the hairpin loop. Simulated annealing and molecular dynamics with 148 distance constraints, 11 hydrogen bonds, and 84 torsion angle constraints showed a wide variety of structures. Certain trends are evident, such as continuation of the A-form helix on the 3′ side of the hairpin loop. The ensemble of calculated structures agree with most chemical modification data.

The expression of the HIV-1 genome in an infected host is regulated by many proteins [e.g., Cullen (1991)]. Early expression is controlled by the positive feedback of the Tat (transcription antitermination) protein, which binds to the TAR (transactivation response) RNA element. The Tat protein interacts with the transcription complex, prevents premature termination of the RNA transcript, and increases the rate of initiation (Kao et al., 1987; Hauber et al.,1987; Frankel, 1992).

The TAR element is a 59-nucleotide hairpin, which contains a three-base bulge and a six-base hairpin loop. Early *in vivo* studies showed that transactivation was influenced by the sequences of these two elements. Feng and Holland (1989) and Berkhout and Jeang (1989) showed that single mutations in the hairpin could decrease transactivation to 8% of wild type *in vivo*, and single-bulge mutations could decrease transactivation to 12% of wild type. An *in vitro* study showed that the Tat protein binds only to the bulge and not to the hairpin (Dingwall et al., 1989). A model of this interaction has been studied by NMR (Puglisi et al., 1992). There is evidence that a cellular protein binds to the hairpin loop (Marciniak et al., 1990; Sheline et al., 1991), and *in vivo* studies suggest that the hairpin is also important for transactivation.

The TAR hairpin loop is only two bases larger than the C[UUCG]G tetraloop previously studied in this laboratory (Cheong et al., 1990; Varani et al., 1991). The strategies developed for the tetraloop were based on complete proton and phosphorus assignment, use of all available backbone torsion angles, and conservative constraints of distances and torsion angles. We used the same strategy in the determination of the TAR hairpin structure.

Recently, two other groups have published NMR studies of the TAR hairpin loop. Michnicka et al. (1993) incorporated ^{13}C- and ^{15}N-labeled nucleotides in a 29-nucleotide fragment of the native TAR hairpin and made few assignments. Colvin et al. (1993) made a 19-base hairpin whose hairpin loop sequence is the same as TAR, but with a different stem and loop closing base pair. Colvin et al. (1993) used a similar strategy to Varani et al. (1991) and obtained an approximate model of the TAR hairpin. The range of structures we obtained by restrained molecular dynamics are compared with the model of Colvin et al. (1993).

MATERIALS AND METHODS

RNA Synthesis and Purification. The oligoribonucleotide 5′-GAGC[CUGGGA]GCUC-3′ ("truncated TAR hairpin") was synthesized enzymatically using T7 RNA polymerase from a synthetic DNA template (Milligan et al., 1987). The hairpin was purified as previously reported (Varani et al., 1989; Puglisi et al., 1990); yields were about 14 nmol of purified RNA per milliliter of reaction. The sequence was confirmed by enzymatic sequencing (Puglisi et al., 1990), and by NMR.

Thermodynamics. The molecularity of the system was determined from UV absorbance versus temperature profiles (Puglisi & Tinoco, 1989). Data were recorded at 260 and 280 nm using a Gilford spectrometer with a thermostated cell holder. The melting curves were independent of the heating rate (0.25, 0.5, and 1 °C/min gave identical melts), and these data were gathered over a 100-fold concentration range (≈ 0.01 to ≈1 mM). All melting samples were dialyzed with 10 mM sodium EDTA, then 1 mM sodium EDTA, and finally 10 mM sodium phosphate, pH 5.5, and 0.2 mM sodium EDTA. Standard enthalpies and entropies were determined from the temperature dependence of the equilibrium constant (Cheong, 1990).

NMR Spectroscopy. NMR samples were dialyzed as described above. Lyophilized samples were dissolved in a 9:1 mixture of H_2O/D_2O to a volume of about 0.4 mL and a concentration of 1–2 mM. Samples for nonexchangeable proton spectra were exchanged at least four times with 99.8% D_2O (Aldrich) in the NMR tube and finally dissolved to 0.4 mL with 99.98% D_2O.

† Supported by NIH GM 10840, NIH GM 14468-02, DOE DE-FG03-86ER60406, and instrumentation Grants DOE DE-FG05-86ER75281, NSF DMB 86-09305 and NSF BBS 86-20134.

* To whom correspondence should be addressed.

‡ Present address: Genta Inc., 3550 General Atomics Ct., San Diego, CA 92121.

® Abstract published in *Advance ACS Abstracts*, October 15, 1993.

0006-2960/93/0432-12522$04.00/0

One-dimensional (1D) proton spectra were recorded on a GE GN-500 (500-MHz proton frequency) or Bruker AMX-600 (600-MHz proton frequency) spectrometer. Exchangeable proton spectra were recorded with the $1\bar{3}3\bar{1}$ pulse sequence (Hore, 1983), and the excitation maximum was set between the imino and aromatic region. The typical temperature for exchangeable NOE experiments was 8 °C. Nonexchangeable proton spectra were recorded with preirradiation of the residual HDO peak. Approximate $T1$ measurements were recorded using a nonselective 180° composite pulse and a relaxation delay of 10 s.

Two-dimensional nonexchangeable spectra were recorded in the phase-sensitive mode using TPPI (Drobny et al., 1979; Marion & Wüthrich, 1983) and preirradiation of the residual HDO peak, at a temperature of 25 °C. All experiments were processed with FTNMR (Hare Research, Inc.).

Nonexchangeable NOESY spectra were recorded on either the GN-500 or AMX-600 spectrometers. Typically, 350 FIDs of 1K complex points were collected with a 5560-Hz sweep width. Experiments with mixing times of 60, 120, and 200 ms were collected sequentially, and zero-quantum coherences were eliminated by incrementing the mixing time with t_1 (Macura & Ernst, 1979). A 400-ms mixing time NOESY was used for assignment only. Repetition delays were typically 2 or 3 s. Nonexchangeable, frequency offset compensated ROESY spectra (Griesinger & Ernst, 1987) were recorded on the AMX-600 spectrometer with a 100-ms mixing pulse and a field strength $\gamma B_1/2\pi = 2.2$ kHz. Acquisition and apodization parameters were the same as for the NOESY experiment. Exchangeable NOESY spectra were recorded on the AMX-600 spectrometer in 90% H_2O/10% D_2O. Spectra were obtained at 5 °C with a $1\bar{1}$ sequence replacing the last pulse (Plateau & Gueron, 1982), with the first maximum either on the aromatic or imino regions. The second pulse of the $1\bar{1}$ sequence was shortened by ≈ 0.1 μs for better solvent suppression. Five hundred FIDs of 2K complex points were collected with a 12 195-Hz sweep width and processed by digital shift subtraction and a 50° squared sine bell in both dimensions.

High-resolution, phosphorus-decoupled double quantum filtered COSY spectra (DQF-COSY) were collected by using the standard pulse sequence with composite mixing pulses (Müller et al., 1986) on the AMX-600, using GARP1 phosphorus decoupling (Shaka et al., 1985). 750 FIDs of 2K complex points of 80 scans each were collected with a 2040-Hz sweep width. Spectra were apodized with 30° shifted sine bells. Final resolution was 1 Hz/point in the t_2 dimension. Two-quantum spectra (2Q) were recorded as proposed by Braunschweiler et al. (1983) on the GN-500 spectrometer. Composite pulses were used for mixing and 180° pulses in the excitation sequence (Müller et al., 1986). The sweep width was 2801 Hz in t_2 and 5602 Hz in t_1. TOCSY (totally correlated spectroscopy) experiments were recorded on the AMX-600 using MLEV-17 decoupling (Bax & Davis, 1985), with a 150-ms mixing time. Acquisition parameters were similar to the NOESY spectra.

Proton-detected ^1H–^{31}P heteronuclear correlated spectroscopy (HETCOR) was recorded as proposed by Sklenar et al. (1986), on the AMX-400 spectrometer (162-MHz phosphorus frequency) with an inverse probe. The experiment was recorded using TPPI; axial peak suppression and CYCLOPS were combined to give an eight-step phase cycle (Varani et al., 1991). A total of 267 FIDs of 2K complex points of 128 scans each were used, with a sweep width of 1400 Hz in the proton dimension and 600 Hz in the phosphorus dimension. The spectrum was apodized as in Sklenar et al. (1986), and the phosphorus frequency was referenced to an external standard of aqueous trimethyl phosphate (TMP).

^1H–^{13}C HMQC (Varani & Tinoco, 1991a) were recorded on the AMX-600 spectrometer (151-MHz carbon frequency), using natural abundance ^{13}C. A low-resolution experiment was acquired with 90 FIDs of 1K complex points of 320 scans, with sweep widths of 5000 Hz in the proton dimension and 20 000 Hz in the carbon dimension. High-resolution experiments were acquired with GARP1 proton decoupling (Shaka et al., 1985) with 90 FIDs of 1K complex points of 320 scans; sweep widths were 2040 Hz in the proton dimension and 6250 Hz in the carbon dimension. The data were apodized and referenced as in Varani and Tinoco (1991a).

Interproton Distances. Distance estimates of nonexchangeable protons were derived from NOESY peak intensities at different mixing times (200, 120, and 60 ms) using the isolated spin pair approximation. Peaks were integrated within FTNMR, and errors were based on volume reproducibility as well as noise levels near the peak. The volumes as a function of time were subjected to a linear weighted fit (Bevington, 1969), and these buildup rates were compared to that of the C4 H5–H6 pair to determine the interproton distances (Wüthrich, 1986). The average distance for all pyrimidine H5–H6 protons was 2.5 ± 0.04 Å, using a reference distance of 2.45 Å for C4 H5–H6. Data from peaks on each side of the diagonal were analyzed separately to obtain distances; the distances were then averaged. In many instances, the 200-ms data showed significant deviation from the other data points; in these cases, the 200 ms point was discarded and the line refit. These distances were converted to the ranges 1.8–3, 2–4, 2.5–5, or > 4 Å for structure determination (Cheong et al., 1990). Strong ROESY peaks were interpreted as interproton distances of 1.8–3.3 Å, as suggested by Bauer et al. (1990).

Exchangeable proton NOEs were given much looser constraints, since relaxation can be due to chemical exchange with water, as well as NOE cross relaxation. Steady-state imino NOEs were classified as very strong (U imino to A H2) or weak (imino to imino) and given constraints of 1.8–3.0 and 3.0–5.0 Å, respectively. Other water NOESY cross peaks were used as minimal constraints of 1.8–5.0 Å. Colvin et al. (1993) report exchangeable NOEs only for assignment purposes.

In cases where the NOE or ROE was to the indistinguishable H5′ or H5″, the distance restraints were applied to the common C5 atom instead. Similarly, amino proton restraints were applied to the common N atom. In either case, the constraint was adjusted to include the length of the chemical bond.

Scalar Coupling Measurements. ^1H–^1H coupling constants were measured from a high-resolution DQF-COSY. Coupling constants were based on averages of active and passive couplings in the t_2 (at 1 Hz/point) dimension. ^1H–^{31}P coupling constants were measured from a high-resolution HETCOR. J_{HP} was determined by subtracting the passive J_{HH} constants for each spin system. Weak or absent couplings in either experiment, such as $J_{1'2'}$ or $J_{H5''P}$, were assumed to be less than the average peak width (3 Hz). Coupling constant error was the larger of the variation from multiple measurements or 3 Hz. Torsion angles were determined from these coupling constants using the generalized Karplus equations of Haasnoot et al. (1980) and Lankhorst (1984). The percentage of C3′-*endo* sugar conformer was determined from $J_{1'2'}$ using the linear equations of de Leeuw and Altona (1982). Torsion angle errors were estimated from errors in the coupling

constants (from multiple measurements) and from the behavior of the generalized Karplus equations at the measured coupling constant values. No proton–proton coupling constants were reported by Michnicka et al. (1993); Colvin et al. (1993) use a 2 Hz/point DQF-COSY (as compared to the 1 Hz/point DQF-COSY reported here) to measure $J_{1'2'}$, and report if the sugar was C2′- or C3′-*endo*. The ^{31}P–^{1}H HETCOR measured by Colvin et al. cannot be used to measure J_{HP} because the detected nucleus (^{31}P) relaxes quickly as compared to ^{1}H; this leads to broad lines along the high-resolution dimension, obscuring all coupling information.

Structure Calculations. The simulated annealing and molecular dynamics program X-PLOR (Brünger, 1992) was used with the CHARM, all-hydrogen, nucleic acid force field (Nisson & Karplus, 1986) for electrostatic and van der Waals interactions, while bond length and angle forces were set to large values [1000 kcal/(mol·Å^2), and 500 kcal/(mol·rad^2), respectively], as suggested by Nilges et al. (1988a). Calculations were carried out in vacuo, with reduced phosphate charges (Tidor et al., 1983). The protocol used is well established (Nilges et al., 1988a; Wimberly et al., 1992; SantaLucia & Turner, 1993). No structural modeling or calculations were reported by Colvin et al. (1993).

Structures with randomized torsion angles were annealed using NOE constraints. The structures were initially annealed at 1000 K with only intranucleotide NOEs, and then annealed with all NOE restraints. This two-step procedure brought all of the structures to the same relative energy minimum; a similar protocol has been used for protein refinement (Nilges et al., 1988b). The annealing protocol used only the repulsive van der Waals term and no electrostatic interactions. The structures were refined with all NOEs and torsion angles β, γ, δ, ϵ, ν_1, and ν_2 at 1000 K; subsequent refinement used all NOE and torsion angle restraints at 1000 K, followed by cooling to 300 K. During refinement the complete van der Waals term was used, and electrostatic forces were evaluated with a reduced phosphate charge of –0.32e (instead of –1e) and a dielectric constant of 1.0 (Tidor et al., 1983). After 3 ps of dynamics at 300 K, the structures were energy minimized.

RESULTS

Thermodynamics. Absorbance melting curves of the molecule showed no change in melting temperature T_m over a 100-fold concentration range, indicating that the system was unimolecular. As the concentration increased, the lower base line increased its slope, perhaps as the result of slight aggregation. Fitting these data (in 10 mM sodium phosphate) to a two-state, unimolecular model yielded $T_m = 65 \pm 0.8$ °C, $\Delta H° = -46 \pm 1$ kcal/mol, and $\Delta S° = -136 \pm 1$ cal/(mol·K). The measured values are within the error of the calculated nearest-neighbor values in 1 M NaCl: $T_m = 64.7$ °C, $\Delta H° = -39.2$ kcal/mol, and $\Delta S° = -116.1$ cal/(mol·K) (Jaeger et al., 1991). The nearest-neighbor thermodynamic model assumes that the truncated TAR hairpin free energy is a sum of six terms: the three nearest neighbors in the stem ($^{GA}_{CU}$, $^{AG}_{UG}$, and $^{GC}_{CG}$), the first and last bases in the loop stacking on the closing base pair ($^{CC}_{G}$ and $^{C}_{GA}$), and the sequence-independent term for closing the hairpin loop ([CUGGGA]).

To test for formation of a C_5·AH^+_{10} base pair the pH dependence of T_m was measured. A pH decrease stabilizes an AH^+·C pair in the [AUUUCUGAC] hairpin; this increases the T_m and allows two more base pairs to form (Puglisi et al., 1990). The truncated TAR hairpin's T_m did not change over a pH range of 4.5–7.5, which could imply that no C_5·AH^+_{10} pair forms. However, thermodynamic calculations of the truncated TAR hairpin (GAGC[CUGGGA]GCUC) and a $C_5 \rightarrow U$ mutant hairpin with a U_5·A_{10} closing base pair (GAGCU[UGGG]AGCUC) gave no appreciable change in T_m when that pair was made; the $C_5 \rightarrow U$ mutant would have an extra nearest neighbor ($^{CU}_{GA}$), two different stacks ($^{UU}_{A}$ and $^{U}_{AG}$), and a four-base loop ([UGGG]). Thus, the lack of a pH dependence of T_m cannot rule out a C_5·AH^+_{10} base pair in the TAR loop.

Spectral Assignments (Table I). Imino assignments were made from temperature behavior and steady-state NOEs in 1D spectra in water at 8 °C. The broad resonance at 12.6 ppm (G1) and the three peaks at 11.05, 10.86, and 10.62 ppm disappear at 35 °C, followed by the 14.16 ppm (U13) resonance and finally the two remaining resonances at 13.51 and 13.32 ppm (G3, G11) (Figure 1). These assignments were confirmed by steady-state NOEs. The A2 H2 (7.60 ppm) was assigned on the basis of the strong steady-state NOE from the U13 imino. The three remaining imino resonances (11.05, 10.86, and 10.62 ppm) required a much higher power for complete saturation and showed no steady-state NOEs. These resonances are in the right range for non-hydrogen-bonded loop iminos (Puglisi et al., 1990; Varani et al., 1991). Colvin et al. (1993) do not see these three broad loop imino resonances in their spectra, even at low pH and low temperature (D. W. Hoffman, personal communication). Assignment of the aminos in the stem and C5 were made with a NOESY experiment at 8 °C, using the assignment pathway of Heus and Pardi (1991b). No loop aminos besides C5 could be assigned. Well-resolved minor form peaks at 14.01, 13.42, 12.42, 11.93, 11.64, and 10.41 ppm were tenatively assigned to the duplex form of the hairpin; these small peaks showed strong cross peaks in the exchangeable NOESY spectrum, suggesting a higher molecular weight species. The ratio of hairpin to duplex was at least 5:1 by peak integration of the U13 imino hairpin and minor form peaks at 0 °C. These minor form peaks disappeared by 25 °C. Addition of 1 mM Mg^{2+} converted most of the hairpin to duplex, as suggested by 1D imino NMR.

Adenine H2 protons usually have characteristically long T_1s (Wüthrich, 1986), although unusual helix geometries [e.g., Chastain and Tinoco (1992)] can produce small AH2 T_1s. Nonselective inversion of protons gives only approximate T_1s (Derome, 1987); nevertheless, these T_1s are useful for AH2 assignment. The resonances at 7.60 and 8.15 ppm have T_1s of approximately 9 and 7 s, as compared to 4 s for the majority of the other aromatic resonances. The resonance at 7.60 ppm corresponds to the A2 H2 resonance found with steady-state NOEs, and the resonance at 8.15 ppm was tentatively assigned to A10 H2.

A DQF-COSY, which assigned all H5–H6 cross peaks, characteristic chemical shifts of 5′ G1 and 3′ C14 nucleotides (Varani & Tinoco, 1991b), and the initial assignments of the two AH2 resonances were used for the aromatic to anomeric proton assignments in a 400-ms NOESY (Figure 2). The stem shows the standard A-form connectivities, both aromatic to H1′ and aromatic to H2′, which continue to A10 in the loop. Weak NOEs are seen to C5, U6, and G9. This pattern of NOE connectivities is similar to those reported by Colvin et al. (1993). The G1 and A2 H8 have similar chemical shifts, but the elongated peak betrays the two resonances. G7, G8, and G9 resonances were assigned from sugar proton connectivities. The G7 H8 to H1′ cross peak is split in the H1′ dimension in both the ROESY and NOESY spectra, suggesting two possible forms. The NOESY spectra show unusual cross peaks between U6 H2′ and G7 H5″, suggesting a sharp

Table I: Proton, Carbon, and Phosphorus Chemical Shifts for the TAR Hairpin[a]

base	H8/H6	H5/H2	H1′	H2′	H3′	H4′	H5′ (H5″)	imino	amino
G1	8.10		5.72	4.89	4.76	4.50	4.29 (4.23)	12.60	
A2	8.10	7.60	6.06	4.76	4.66	4.56	4.47 (4.33)		6.58/8.00
G3	7.22		5.62	4.43	4.37	4.45	4.41 (4.12)	13.54	6.18/8.65
C4	7.44	5.12	5.44	4.27	4.52	4.37	4.54 (4.04)		7.00/8.50
C5	7.68	5.57	5.82	4.32	4.50	4.22	4.44 (3.97)		7.35/7.78
U6	7.68	5.70	5.57	4.21	4.45	3.99	4.44 (4.07)	*b*	
G7	7.78		5.47	4.61	4.54	4.08	4.02 (3.93)	*b*	
G8	7.64		5.51	4.51	4.63	4.02	3.83 (3.79)	*b*	
G9	7.88		5.90	4.90	4.99	4.47	4.43 (4.13)	*b*	
A10	8.42	8.15	5.89	4.85	4.74	4.57	4.33 (4.32)		
G11	7.36		5.50	4.53	4.41	4.46	4.20 (4.14)	13.35	6.94/7.65
C12	7.71	5.24	5.51	4.34	4.45	4.41	4.52 (4.03)		6.97/8.60
U13	7.89	5.45	5.58	4.28	4.52	4.34	4.53 (4.02)	14.20	
C14	7.73	5.70	5.83	3.96	4.17	4.14	4.44 (4.00)		7.16/8.27

base	C8/C6	C5/C2	C1′	C2′	C3′	C4′	C5′	P[c]
G1	142.5		92.4	78.2/77.6	78.1/79.0	86.6	70.0	
A2	142.5	155.9	95.2	78.1	75.9	87.1/85.2	69.2	−3.50
G3	138.3		95.3	78.6	75.3	≈84.5	68.5	−3.88
C4	143.1	99.6	96.9	77.9	74.5/77.4	≈84.6	66.7	−4.38
C5	145.4	101.3	95.0	74.4	74.5/77.4	87.1	67.7	−4.12
U6	145.4	107.6	96.6	78.1	74.5	87.3	68.2	−4.12
G7	143.1		91.9	77.4	80.2	87.6	70.0	−3.76
G8	142.2		95.8	78.4	79.8	87.3	70.1	−3.80
G9	142.2		92.4	78.2/77.6	78.1/82.0	87.0	68.5	−3.81
A10	145.6	154.8	92.4	78.2	78.1	87.1/85.2	70.6	−3.40
G11	139.3		92.1	77.7	74.5/77.4	85.2	70.2	−3.58
C12	144.3	102.4	96.7	78.0	74.5	≈84.4	66.7	−4.63
U13	144.7	105.9	93.9	77.9	74.5/77.4	85.9/≈84.5	66.7	−3.76
C14	143.9	100.6	95.0	82.9	72.4	85.9	67.8	−3.98

[a] All proton chemical shifts referenced to 3-(trimethylsilyl)propionate (TSP), at 10 mM sodium phosphate, pH 5.5. Imino and amino resonances were measured at 8 °C. All other proton, carbon, and phosphorus resonances were measured at 25 °C. [b] Three imino resonances at 11.05, 10.86 and 10.62 could not be assigned. [c] Phosphorus chemical shifts relative to trimethyl phosphate (TMP).

Table II: Proton–Proton and Proton–Phosphorus Coupling Constants for the TAR Hairpin

base	$J_{1'2'}$	$J_{2'3'}$	$J_{3'4'}$	$J_{4'5'}$ ($J_{4'5''}$)	$J_{4'Pi}$	$J_{5'Pi}$ ($J_{5''Pi}$)	$J_{3'P(i+1)}$
G1	4	4	6	<3 (5)			9
A2	3	5	10	<3 (6)	<3	2–5 (<3)	7
G3	<3	5	10	<3 (6)	<3	<3 (<3)	6
C4	<3	5	10	<3 (5)	<3	<3 (<3)	11
C5	4	5	9	<3 (8)	<3	3–6 (<3)	12
U6	<3	5	10	<3 (<3)	<3	<3 (<3)	6
G7	8	4	7	<3 (7)	<3	<3 (3.5)	5
G8	6	4	10	<3 (6)	<3	11–14 (8)	3
G9	4	6	10	<3 (5)	<3	4–7 (<3)	3
A10	<3	6	12	<3 (6)	7	<3 (7)	3
G11	<3	5	10	5 (5)	5	<3 (<3)	6
C12	<3	5	10	<3 (5)	<3	3–6 (<3)	3
U13	<3	5	10	<3 (6)	<3	5–8 (<3)	5
C14	<3	5	9	4 (4)	9	<3 (6)	

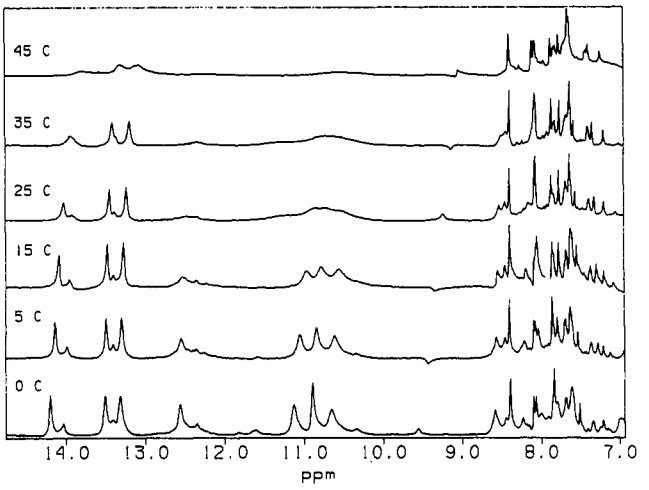

FIGURE 1: H_2O 1D melts in 10 mM sodium phosphate, pH 5.5, and 0.2 mM sodium EDTA. The temperature is listed just above the spectra. The downfield stem resonances are U13, G3, G4, and G1 (broad), at 14.20, 13.54, 13.35, and 12.60 ppm. The three broad resonances between 12.5 and 10.5 ppm could not be assigned, but are in the right region for hairpin loop iminos. These exchangeable proton spectra were recorded with the $1\bar{3}3\bar{1}$ pulse sequence (Hore, 1983), with the excitation maximum set between the imino and aromatic region. The peak between 9 and 10 ppm is a spectrometer artifact.

bend at this point in the hairpin. This crosspeak was not reported by Colvin et al. (1993), since they did not assign any H5′ or H5″ resonances. G7 H4′ has a weak cross peak to G8 H8. Unfortunately, overlap in the spectra prevented any observation of a G8 H1′ or H2′ to G9 H8. G9 H2′ shows a strong cross peak to A10 H4′, completing the walk around the hairpin loop. The A2 H2 resonance clearly shows the cross-strand walk from G3 H1′ to C14 H1′, but the A10 H2 resonance shows only one A-form cross peak to G11 H1′; these NOEs are identical to those of Colvin et al. (1993). No additional A10 H2 cross peaks were seen by increasing the repetition delay to 6.5 s with a 400-ms mixing time, or lowering the temperature to 8 °C. These results suggest the loop bases C5, U6, and A10 are not in a standard A-form orientation with respect to each other.

Correlated spectra were the most important means of spectral assignment for the remainder of the sugar protons (Varani & Tinoco, 1991a). The DQF-COSY provides the assignments for many of the protons, except for H1′ in C3′-*endo* sugars. The 2Q experiment helps resolve cross peaks close to the diagonal, such as the almost overlapping G7 H5′ and H5″. The ^1H–^{31}P HETCOR determines the sequence of the sugars along the backbone; the use of a lower field spectrometer helps resolve ^{31}P resonances, and thus all H3′(i) to H5′ or H5″(i+1) connectivities. Finally, the ^1H–^{13}C (natural abundance) HMQC (Varani & Tinoco, 1991a)

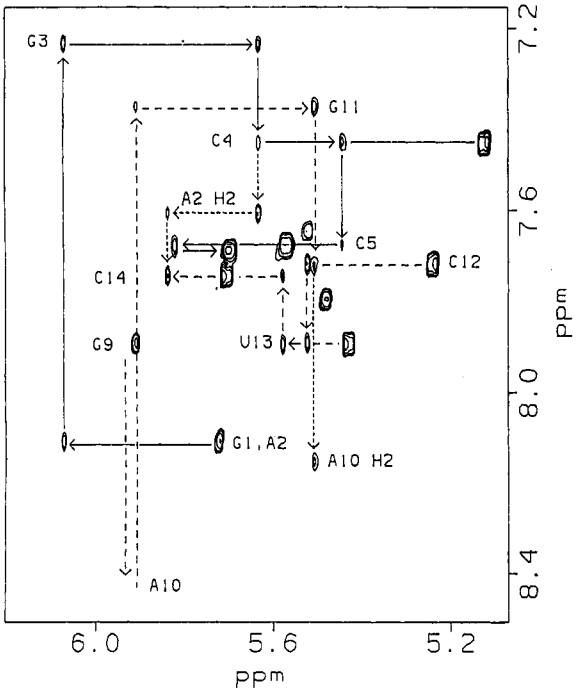

FIGURE 2: A 400-ms NOESY fingerprint (aromatic to anomeric) region acquired at 25 °C, in 10 mM sodium phosphate, pH 5.5, and 0.2 mM sodium EDTA. The walk 5′-GAGC[CU (stem bases G1–C4, and hairpin loop bases C5 and U6) is shown with a solid line, while the walk GA]GCUC-3′ (hairpin loop bases G9 and A10, and stem bases G11–C14) is shown with a long-dashed line. Connectivities to pyrimidine H5–H6 cross peaks are also shown. AH2 walks are shown with short-dashed lines. Arrowheads at the ends of the connectivities point from 5′ to 3′. The spectrum is labeled with the same numbering as in Figure 4. Each pyrimidine H6 has cross peaks to the previous H1′, its own H1′, and its H5; these interactions form three horizontal cross peaks in a row. The H6–H5 cross peak usually shows no additional vertical connectivities. G3 H1′ has cross peaks to G11 H8, C12 H6, and A10 H2; these interactions form three vertical cross peaks in a column. Since G9 H1′ and A10 H1′ are overlapped, the G9 to A10 connectivity is offset from the A10 to G11 connectivity. The G9 H1′–A10 H8 cross peak is present, but weak.

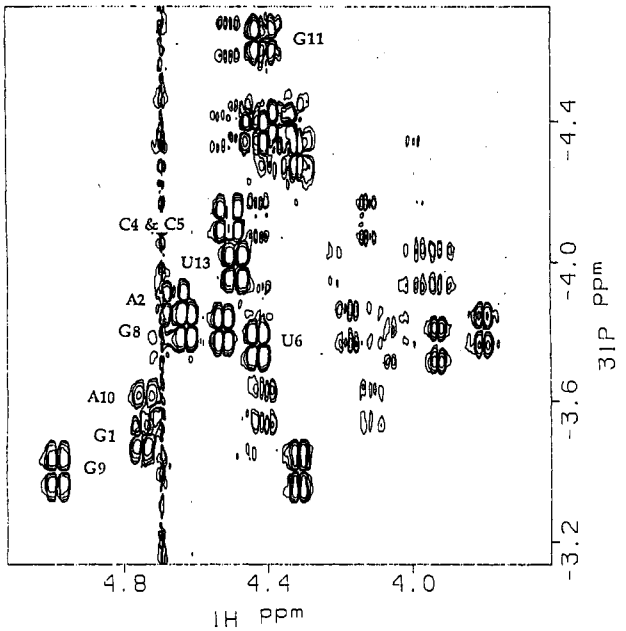

FIGURE 3: Proton–phosphorus HETCOR spectra of the truncated TAR RNA hairpin in 10 mM sodium phosphate, pH 5.5, and 0.2 mM sodium EDTA, at 25 °C. The phosphorus frequency was referenced to an external standard of aqueous trimethyl phosphate (TMP). Peaks corresponding to the internucleotide H3′(*i*-1)–P(*i*) connectivities are labeled with the nucleotide containing the H3′-(*i*-1).

confirmed the proton assignment. The stereospecific assignments of H5′ and H5″ resonances reported are based on chemical shift; the reported H5′ resonance resonates downfield from the H5″, as in normal A-form helices (Varani et al., 1991). As mentioned above, the assignment assumptions of H5′ and H5″ were not used in modeling.

The loop proton assignments of Colvin et al. (1993) are almost identical to ours, but Colvin et al. (1993) do not report assignments for H5′ or H5″ protons. The assignments of G9 H8 and A10 H8 (this paper's numbering scheme) reported by Colvin et al. are swapped relative to those reported here. The assignment of the A10 H8 by Michnika et al. (1993) is almost identical to the one reported here. Both Michnicka et al. (1993) and Colvin et al. (1993) use ^{1}H–^{13}C HMQC and report some carbon assignments.

Backbone Coupling Constants and Torsion Angles (Table II). Torsion angles form important constraints where NOESY data are ambiguous or absent; the torsion angle nomenclature of Altona (1982) is used in this paper.

The α and ζ backbone angles were estimated from phosphorus chemical shift, as suggested by numerous studies of tRNA and DNA (Gorenstein & Luxon, 1979; Salemink et al., 1979; Gorenstein, 1984). The dispersion of the phosphorus resonances in the truncated TAR hairpin is over a 1 ppm range, suggesting α and ζ angles are *gauche* (Gorenstein, 1984). Both α and ζ were constrained to ±100° (Varani et al., 1991).

Coupling constants for β and ϵ are measured from the ^{1}H–^{31}P HETCOR (Figure 3). In standard A-form RNA, β is *trans* (≈180°). Stereospecific assignment of H5′ and H5″ cannot be made without structural assumptions (H5′ resonates downfield of H5″ in A-form RNA) or carbon coupling data. Usually coupling was only seen to the downfield proton. β's error limits were typically ± 25°, except for G8, A10, G11, and C14. A10, G11, and C14 show an additional P(*i*)–H4(*i*) coupling, which only occurs for the *trans* β and *gauche*$^+$ γ conformer; this reduced these nucleotides' β error to ± 20°. Finally, G8 has a large coupling constant for both protons, suggesting β is *gauche* rather than *trans*. This suggests a sharp bend in the helix between G7 and G8. Torsion angle ϵ was always near the standard A-form conformation (*gauche*$^-$ or ≈−120°), as measured by the H3′(*i*)–P(*i*+1) couplings, with typical error limits of ±35°.

The error used for γ is large, since there are no stereospecific H5′/H5″ assignments. This angle was estimated as 60 ± 40° for most bases, consistent with the standard A-form value. A10, G11, and C14 have visible P(*i*)–H4(*i*) coupling; this further restrains γ to 50 ± 30° for these nucleotides.

The ribose ring was constrained by three angles: ν_1, ν_2, and δ (or ν_3). In most cases, the sugars were C3′-*endo*, as shown by the small $J_{1'2'}$ (typically less than 3 Hz) and the large $J_{3'4'}$ (≈9 Hz) couplings. This combination of measurements limited the ribose angles to small ranges. Similarly, the very large $J_{1'2'}$ and small $J_{3'4'}$ of G7 indicates a C2′-*endo* conformation. G1, C5, G8, G9, and C14 all show large $J_{1'2'}$ and $J_{3'4'}$, which could only occur in dynamic sugars, or in an O4′-*endo* conformation. The latter was ruled out by the ≈3-Å H1′–H4′ distance of these sugars; in C2′- or C3′-*endo* sugars, this distance is ≥3 Å, and in O4′-*endo* sugars it is ≈2.2 Å (Wüthrich, 1986). The percent of C2′-*endo* conformation for G1, C5, G8, G9, and C14 was respectively 50%, 40%, 70%, 40%, and < 20%, using the equation % C2′-*endo* =

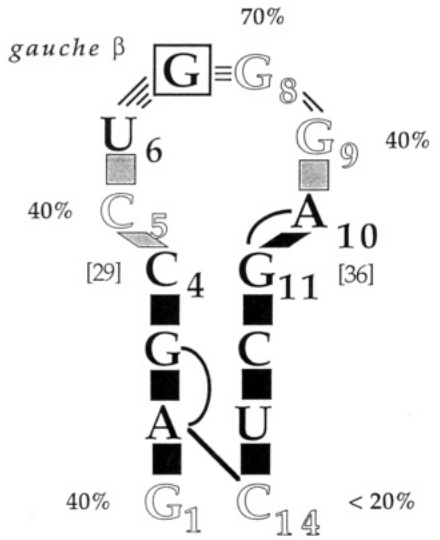

FIGURE 4: A schematic summarizing the NMR data in this study for the truncated TAR hairpin. Sequence numbers in brackets are for comparison with Michnika et al. (1993) and others. The number of internucleotide NOEs between U6, G7, G8, and G9 are represented by the number of lines. The long-range AH2 to H1 NOEs are shown as thick lines. Sugar pucker is either C3′-*endo* (plain text), C2′-*endo* (boxed text), or in equilibrium (outline text, with % C2′-*endo* next to base). A-form NOEs are shown by black boxes between bases, and weak A-form connectivities are shown by shaded boxes. The unusual G7 *gauche* β is marked; the other βs in the structure are the standard A-form RNA *trans* conformation.

$14J_{1'2'}$ - 13 (de Leeuw & Altona, 1982). ν_1, ν_2, and δ for these bases was limited to an angle range encompassing both C3′- and C2′-*endo* conformers.

Distance Ranges. NOESY buildup rates were compared to the rate of buildup of H5–H6 of cytosine 4 to determine interproton distances (Wüthrich, 1986). The buildup rates for all pyrimidine H5–H6 were similar; the mean distance was 2.50 Å, as compared to the C4 H5–H6 reference distance of 2.45 Å. All distances for the helix were within normal A-form range. None of the bases were in a *syn* conformation, as shown by NOESY and ROESY spectra of the aromatic to H1′ region.

Hydrogen bonds in the stem were set to be equal to the ranges given by Saenger (1983). The O–H distance in O·HN hydrogen bonds was constrained to 1.83–2.17 Å; the N–H distance in N·HN hydrogen bonds was constrained to 1.78–2.02 Å.

The NOE and coupling information from this paper is schematically shown in Figure 4.

STRUCTURE ANALYSIS

Fifty structures with random torsion angles were refined as described in Materials and Methods. Initial X-PLOR calculations of the truncated TAR hairpin showed that the structures collapsed into the major groove, and many structures had close contact of A10 H2 with protons of C4, C5, and U6. These initial structural problems were corrected by finding ^{1}H–^{1}H contacts in the calculated structures less than 5 Å, but not suggested by NOE or ROE cross peaks; 15 constraints were set to a range of 4.5–13.9 Å. A similar method was used by White et al. (1992) for an X-PLOR calculation of an RNA helix. Not all structures completely minimized; total energies ranged from −188 to +354 kcal/mol. The 10 structures with the lowest total energy (between −188 and −143 kcal/mol) were studied in depth.

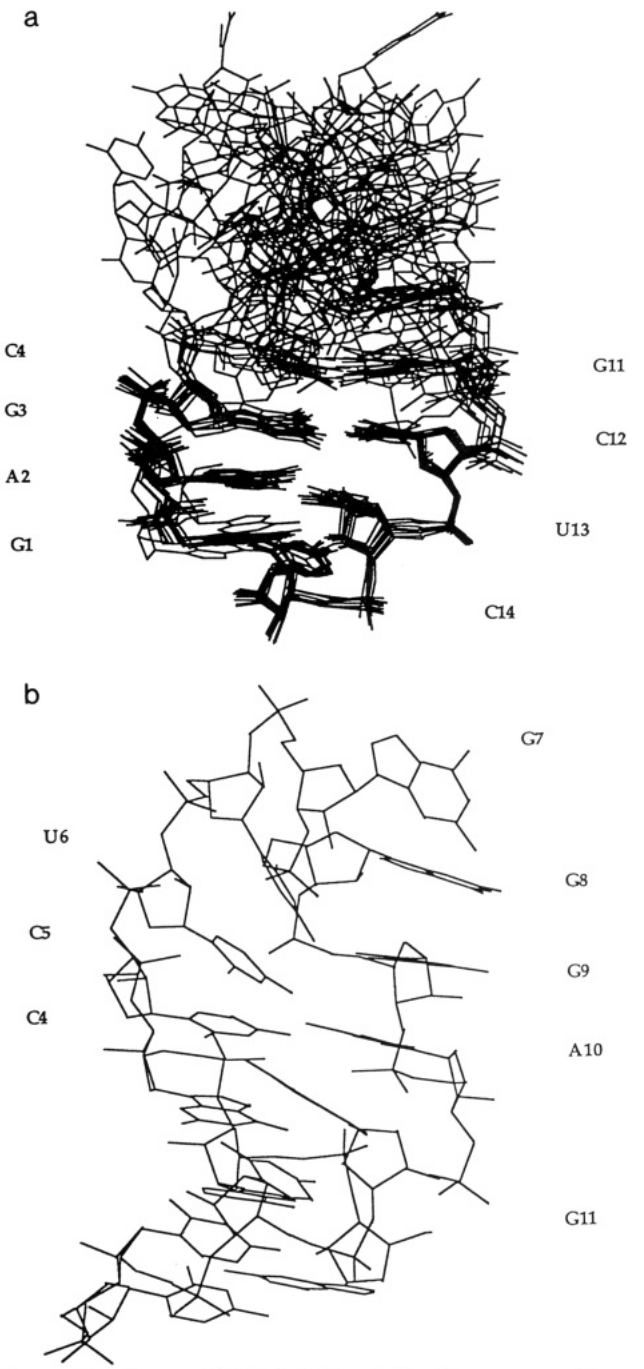

FIGURE 5: Truncated TAR hairpin calculated structures with no extra hydrogen bonds in the hairpin loop. (a) Overlap of ten best truncated TAR hairpin structures, with helix nucleotides numbered. (b) One of the truncated TAR hairpin structures showing stack continuation on the 3′ side of the loop. Discernible bases are numbered. This view is slightly different from that in part a.

Stem. These structures' helices are similar (Figure 5a) and have an average RMSD of 0.64 Å. The central two base pairs of the stem are better defined, as they are constrained by the paired bases at the end of the helix; the RMSD for the two central base pairs is 0.37 Å. C_4·G_{11} is much less defined than the terminal G_1·C_{14}; presumably this is due to flexibility in the loop region. The overall RMSD for the top 10 structures is 2.85 Å.

Hairpin Loop. The hairpin loop region shows a great deal of flexibility (Figure 5a). This was first suggested by the equilibrium states of the C5, G8, and G9 sugars at 40%, 70%, and 40% C2′-*endo*, respectively. The calculated loop variability was chiefly due to the switching of α and ζ between

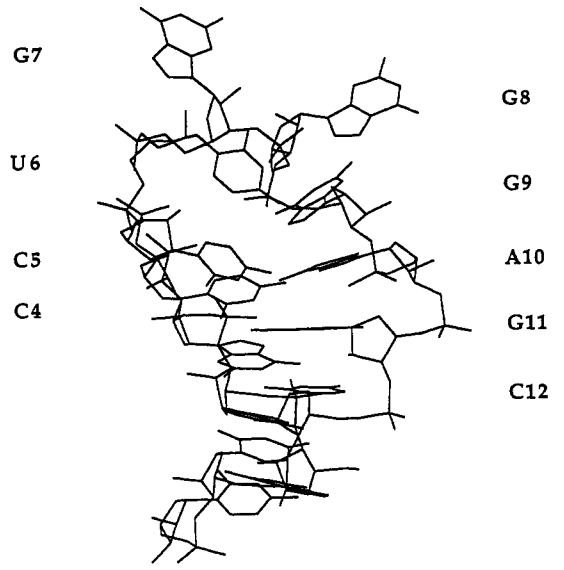

FIGURE 6: Truncated TAR hairpin structures modeled with four hydrogen bonds in the hairpin loop (C5 N3·A10 H6, C5 H4·A10 N7, U6 H3·G9 O6, U6 O4·G9 H1), and all available NOE and torsion angle constraints.

g^- (standard A-form) and g^+. Another variable angle in the loop is G8 β, which is either g^- or g^+ (the standard A-form value is t). Thus there are some 2^{15} = 32 768 possible conformations based on these 15 angles alone (14 α and ζ in the loop and 1 variable β). There is no NMR evidence of extra hydrogen bonds to constrain these loop angles, as was the case in the C[UNCG]G (Varani et al., 1991) and G[GNRA]C (Heus & Pardi, 1991a) hairpin loops. As mentioned before, α and ζ can only be estimated from phosphorus chemical shifts.

All ten structures have A10 stacked on the helix; in seven cases G9 stacks on A10, and in six cases G8 continues the A-form stack. Calculations of structures with only the experimental stem constraints (and no loop constraints) showed no trend of base stacking on the 3′ side of the loop; thus the continuation of the A-form stack is driven by the constraints, not the X-PLOR potentials. RNA hairpin loop base stacking on the 3′ side of the loop was suggested by Haasnoot et al. (1986), based on geometrical arguments. They showed that the shortest distance between phosphates in an RNA hairpin occurs when five bases form an A-form stack; this 13-Å distance can be spanned by one or two nucleotides. The only other example of stacking on the 3′ side of an RNA hairpin loop with more than four bases is in the anticodon loop of $tRNA^{Phe}$, but that loop has significant intermolecular contacts in the crystal [i.e., Sussman et al. (1978)]. C5 does not stack on C4, but is in a similar position for all ten structures. U6 and G7 have different positions for all ten structures.

MODELING

A more structured loop was modeled by adding additional hydrogen bonds to the parameter set. There are many possible ways that nucleic acid bases can form pairs (Saenger, 1983); some of the pairs that could form in the TAR hairpin are discussed here. As before, the close (less than 5 Å) ^1H–^1H contacts found in the calculated structures, but not suggested by NOE or ROE cross peaks, were eliminated by constraining them to a range of 4.5–13.9 Å.

C_5·A_{10} and U_6·G_9 could continue the A-form helix, like the U·G pair in the C[UUCG]G tetraloop (Varani et al., 1991). A C·AH^+ pair has been found at pH 5.5 in a hairpin structure, sandwiched between two Watson–Crick pairing regions (Puglisi et al., 1990). This pair was composed of one A N6H·C N3 hydrogen bond and a protonated A $N1H^+$·C O2 hydrogen bond. At pH 5.5, the C·AH^+ base pair plus two other base pairs (A·U, G·U) formed, and the T_m of the hairpin increased. Similar pH studies carried out on the truncated TAR hairpin showed no change in T_m with pH, but in the TAR loop only one new base pair can form. Nearest-neighbor thermodynamic calculations of the truncated TAR hairpin ([CUGGGA]) and a $C_5 \rightarrow U$ mutant hairpin resulting in a U·A closing base pair (U[UGGG]A) suggest no appreciable change in T_m should occur when that pair was made. However, both C_5 and A_{10} show many A-form NOEs, suggesting that the stem may extend by this new base pair. This base pair was modeled as a single C_5N3·A_{10}N6H pair, since the other hydrogen bond could be formed by protonation, or a bound water. The structures calculated with this single C·A hydrogen bond were consistent with the NMR data. The extra hydrogen bond decreased the stacking on the 3′ side of the loop to four out of ten structures for G9, and none for G8. C5 does not stack on C4, and its base is in a different position in each structure. A U_6·G_9 pair would protect two iminos in the loop and constrain the hairpin loop to only two nucleotides. A similar geometry is seen in the C[UUCG]G hairpin (Varani et al., 1991), where a severely bent U·G pair closes a two-base hairpin loop, albeit with a *syn* G. In the truncated TAR hairpin, there are no characteristic U·G resonances, and none of the bases appear *syn*. Nevertheless, the NMR data are consistent with both U_6·G_9 and C_5·A_{10} pairs in the same molecule (Figure 6). These constraints also prevent stacking on the 3′ side of the loop beyond A10. Similar results were obtained for combinations of the other suggested C·A pairs (C N3·A N6H and C N4H·A N7 or C N3·A N6H and C N4H·A N1) or U·G pair (UN3H·GO6 and U O4·G N1H) listed by Saenger (1988).

The single Watson–Crick pairs C_5·G_9 or U_6·A_{10} could also form. The U_6·A_{10} and C_5·G_9 models show at best strained base pairs in all ten structures. Even though they are consistent with the NMR data, the hydrogen-bonded bases in these systems are usually not coplanar, since the other NOE and torsion angle constraints tend to keep A_{10} and G_9 stacked on the helix.

DISCUSSION

The TAR hairpin is important in the transactivation of HIV-1 transcription *in vivo*, and its position and length are conserved (Berkhout, 1992). Mutations in the hairpin can greatly diminish transactivation in vivo. Unfortunately, there are many candidates for the cellular protein that interacts with this hairpin (Marciniak et al., 1990; Sheline et al., 1991). What does the solution structure tell us about hairpin structure in general, and about the cellular protein that recognizes the TAR hairpin?

This NMR study suggests that this six-base hairpin loop is very flexible in solution, as shown by the large $J_{1'2'}$ and $J_{3'4'}$ couplings of C5, G8, and G9. The trend of previous hairpin loop structures suggests this loop would form the C5·A10H^+ and U6·G9 pairs and have a strained loop of the two remaining Gs. What is different about this loop? In the previous study of a hairpin loop with a C·AH^+ pair, the nine-base loop [AUUUCUGAC] formed three base pairs and a three base loop AUU[UCU]GAC in solution (Puglisi et al., 1990). The C·AH^+ pair may be more likely to form in the context of surrounding stable helical regions. A similar context-dependent effect is seen for the thermodynamics of terminal versus internal $\begin{smallmatrix}GA\\AG\end{smallmatrix}$ pairs (SantaLucia et al., 1991), and to a lesser

Table III: Comparison of Calculated Probe Accessibility and Chemical Modification Data for the TAR Hairpin Loop

position	probe	rel modification intensity	model probe radius (Å)[e]	av accessibility (Å^2)
C5 N3	DMS	none[a]	1.8	0
U6 N3	CMCT	strong[a]	3.0	0
G7 N1	kethoxal, CMCT	strong[a,c]	3.0	0
G7 N7	DMS	*d*	1.8	13
G8 N1	kethoxal, CMCT	medium[a,c]	3.0	0
G8 N7	DMS	*d*	1.8	9
G9 N1	kethoxal, CMCT	weak[a,c]	3.0	0
G9 N7	DMS	*d*	1.8	5
A10 N1	DMS	strong[a]	1.8	8
A10 N7	DEPC	strong[b]	3.0	0

[a] Colvin & Garcia-Blanco, 1992. [b] Berkhout, 1992. [c] G7, G8, and G9 show similar reactivity in the N2 position (data not shown) for kethoxal modification. [d] No data reported in the literature. [e] Lavery & Pullman, 1984.

extent terminal versus internal G·U pairs (Freier et al., 1986; Sugimoto et al., 1986).

The TAR hairpin loop structure follows the rules originally proposed by Haasnoot et al. (1984), based on the crystal structure of tRNA^{Phe}; there is more base stacking on the 3′ side of the hairpin loop, and less on the 5′ side of the loop. Stacking on the 3′ side of the loop is opposite the trend of RNA thermodynamic measurements of dangling ends, which show a large contribution of 3′ dangling ends (corresponding to the 5′ side of the loop) and a negligible contribution of 5′ dangling ends (corresponding to the 3′ side of the loop). This discrepancy is probably due to the environment since the hairpin loop nucleotides are more constrained than the simple dangling nucleotide.

The flexible TAR loop model can be independently checked by comparing the base surface accessibility with experimental chemical modification data. The four probes, DMS (dimethyl sulfate), DEPC (diethyl pyrocarbonate), CMCT [1-cyclohexyl-3-(2-morpholinoethyl)carbodiimide methyl *p*-toluenesufonate], and kethoxal (3-ethoxy-1,1-dihydroxy-2-butanone), were modeled as spheres rolling over the hairpin's van der Waals surface (Holbrook & Kim, 1983) with radii of 1.8, 3.0, 3.0, and 3.0 Å, respectively (Lavery & Pullman, 1984), using the X-PLOR surface function. An average accessibility was calculated for all ten structures. There is very good agreement with the kethoxal (N1 and N2 atoms) and CMCT (N1) modification patterns of G7, G8, and G9, DMS modification of A10's N1 (Colvin & Garcia-Blanco, 1992), and A10's N7 modification by DEPC (Berkhout, 1992) (Table III). The strong modifications of U6's N3 by CMCT (Colvin & Garcia-Blanco, 1992) is not suggested by these models. Predicted loop G N7 accessibility by DMS is also listed in Table III, but no modification data have been reported (unlike DMS modification of U N3 or A N1, G N7 modifications require a subsequent reaction with aniline for detection). Both Holbrook and Kim (1983) and Lavery and Pullman (1984) see no good correlation of DEPC modification of tRNA with spheres of any radius.

A similar analysis of imino proton accessibility, using a 1.6-Å radius sphere to represent a water molecule, shows very similar accessibilities of the four loop imino protons. As a group, these structures do not explain why three broad imino resonances are still visible in 1D water spectra at 25 °C, pH 5.5, in 10 mM sodium phosphate. Individual structures show protection of one or more iminos. An alternate explanation for decreased imino proton accessibility is additional structure.

What structural clues might a cellular protein that interacts with the TAR loop recognize? Mutation studies show that the most important bases in the loop in order of transactivation efficiencies are G7 > C5 > G9 > G8 > U6 > A10 (Berkhout & Jeang, 1989). This apparent recognition of the 5′ side of the loop is contrary to the recognition of the anticodon loop of tRNA^{Gln} by its synthetase (Rould et al., 1991) and the suggested interaction of the U1 A protein with loop II of U1 snRNA (Nagai et al., 1990). The most likely recognition element on the 5′ side of the loop (after sequence) is the series of a C3′-*endo* sugar of U6, followed by the *gauche* β and the C2′-*endo* sugar of G7.

The work of Colvin et al. (1993) shows significant differences with our work, while the limited data of Michnika et al. (1993) agree with our assignments. There are four notable differences between Colvin et al. (1993) and this work. The assignments of the G9 and A10 H8 nuclei are interchanged compared to this molecule. The one assignment by Michnika et al. of the aromatic C and H nuclei of A10 agrees with the data presented here. The T_1 lifetime of the A10 H2 proton reported by Colvin et al. is much shorter than reported here. Colvin et al. see no extra resonances corresponding to the four loop iminos; we report three unassigned, broad resonances in the characteristic region for non-base-paired iminos. Finally, all of the sugar puckers in the loop of Colvin et al. are C2′-*endo*, whereas the data here clearly show sugars ranging from C3′-*endo* , to equilibrium, to C2′-*endo*. These data suggest that the structure of the hairpin loop of Colvin et al. is different from that of Michnika et al. and the one reported here; the hairpin loop of Colvin et al. is closed by a G·C pair rather than the conserved C·G pair (this work and Michnika et al.), which may account for some of the differences.

CONCLUSIONS

The truncated TAR hairpin proton spectrum has been completely assigned, except for the exchangeable protons in the loop and the stereospecific assignments of H5′ and H5″. Homonuclear NMR provided all of the assignments and distance measurements, and many of the torsion angle measurements (ν_1, ν_2, and ν_3 or δ, γ, and the % C2′-*endo*). ^{1}H–^{31}P HETCOR gave phosphorus assignments, and the other torsion angles (β, ϵ) and natural abundance ^{1}H–^{13}C HMQC confirmed the proton assignments. Random torsion angle structures were the starting point for simulated annealing and restrained molecular dynamics using X-PLOR.

There is no one structure for this hairpin in solution, but there are structural trends. Many of the calculated structures show stacking on the 3′ side of the stem. There is no compelling evidence that the loop is closed by any non-Watson–Crick pairs, such as a C·AH$^+$ or G·U base pair. However, formation of these hydrogen bonds is allowed by the NMR data. The phosphorus resonances for this hairpin are all within 1 ppm of each other, suggesting that none of the α or ζ torsional angles are *trans*; other structures of RNA hairpins all have at least one α or ζ *trans*. The truncated TAR hairpin shows unusual NOEs and a *gauche* G7 β, the result of a sharp bend in the loop. While 1D NMR water spectra at low temperature (0 °C) and low pH (5) show three broad but resolved resonances in the loop imino region, these iminos could not be assigned. These peaks suggest that the loop has a more compact structure than we are able to determine from the NMR data.

ACKNOWLEDGMENT

We would like to thank Mr. David Koh for synthesis of DNA templates and Ms. Barbara Dengler for managing the

laboratory. In addition, we would like to thank Dr. Brian Wimberly, Dr. Michael Chastain, and Dr. Gabriel Varani for helpful discussions and Dr. Joseph Puglisi and Ms. Suman Mirmira for critically reading the manuscript.

REFERENCES

Altona, C. (1982) *J. R. Neth. Chem. Soc. 101*, 413–433.

Bauer, C. J., Frenkiel, T. A., & Lane, A. N. (1990) *J. Magn. Reson. 87*, 144–152.

Bax, A., & Davis, D. G. (1985) *J. Magn. Reson. 65*, 355–360.

Berkhout, B. (1992) *Nucleic Acids Res. 20*, 27–31.

Berkhout, B., & Jeang, K.-T. (1989) *J. Virol. 63*, 5501–5504.

Bevington, P. R. (1969) *Data Reduction and Error Analysis for the Physical Sciences*, McGraw Hill, New York.

Braunschweiler, L., Bodenhausen, G., & Ernst, R. R. (1983) *Mol. Phys. 48*, 535–560.

Brünger, A. T. (1992) *X-PLOR Version 3.0 User Manual*, Yale University, New Haven, CT.

Chastain, M., & Tinoco, I., Jr. (1992) *Biochemistry 31*, 12733–12741.

Cheong, C. (1990) Thesis, University of California, Berkeley.

Cheong, C., Varani, G., & Tinoco, I., Jr. (1990) *Nature (London) 346*, 680–682.

Colvin, R. A., & Garcia-Blanco, M. A. (1992) *J. Virol. 66*, 930–935.

Colvin, R. A., White, S. W., Garcia-Blanco, M. A., & Hoffman, D. W. (1993) *Biochemistry 32*, 1105–1112.

Cullen, B. R. (1991) *FASEB J. 5*, 2361–2368.

de Leeuw, F. A. A. M., & Altona, C. (1982) *J. Chem. Soc., Perkin Trans. 2* , 375–384.

Derome, A. E. (1987) *Modern NMR Techniques for Chemistry Research*, Pergamon Press, New York.

Dingwall, C., Ernberg, I., Gait, M. J., Green, S. M., Heaphy, S., Karn, J., Lowe, A. D., Singh, M., Skinner, M. A., & Valerio, R. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 6925–6929.

Drobny, G., Pines, A., Sinton, S., Weitekamp, D. P., & Wemmer, D. (1979) *Faraday Symp. Chem. Soc. 13*, 49–55.

Feng, S., & Holland, E. C. (1988) *Nature (London) 334*, 165–167.

Frankel, A. D. (1992) *Curr. Opin. Genet. Dev. 2*, 293–298.

Freier, S. M., Kierzek, R., Caruthers, M. H., Neilson, T., & Turner, D. H. (1986) *Biochemistry 25*, 3209–3213.

Gorenstein, D. G. (1984) *Phosphorus-31 NMR: Principles and Applications* (Gorenstein, D. G., Ed.) Academic Press, New York.

Gorenstein, D. G., & Luxon, B. A. (1979) *Biochemistry 18*, 3796–3804.

Griesinger, C., & Ernst, R. R. (1987) *J. Magn. Reson. 75*, 261–271.

Haasnoot, C. A. G., de Leeuw, F. A. A. M., & Altona, C. (1980) *Tetrahedron 36*, 2783–2792.

Haasnoot, C. A. G., Hilbers, C. W., van der Marel, G. A., van Boom, J. H., Singh, U. C., Pattabiraman, N., & Kollman, P. A. (1986) *J. Biomol. Struct. Dyn. 3*, 843–857.

Hauber, J., Perkins, A., Heimer, E. P., & Cullen, B. R. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 6364–6368.

Heus, H. A., & Pardi, A. (1991a) *Science 253*, 191–194.

Heus, H. A., & Pardi, A. (1991b) *J. Am. Chem. Soc. 113*, 4360–4361.

Holbrook, S. R., & Kim, S.-H. (1983) *Biopolymers 22*, 1145–1166.

Hore, P. J. (1983) *J. Magn. Reson. 55*, 283–300.

Jaeger, J. A., Turner, D. H., & Zuker, M. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 7706-7710.

Kao, S.-Y., Calman, A. F., Luciw, P. A., & Peterlin, B. M. (1987) *Nature (London) 330*, 489–493.

Lankhorst, P. P., Haasnoot, C. A. G., Erkelens, C., & Altona, C. (1984) *J. Biomol. Struct. Dyn. 1*, 1387–1405.

Lavery, R., & Pullman, A. (1984) *Biophys. Chem. 19*, 171–181.

Macura, S., & Ernst, R. R. (1979) *Mol. Phys. 41*, 95–117.

Marciniak, R. A., Garcia-Blanco, M. A., & Sharp, P. A. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 3624–3628.

Marion, D., & Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun. 113*, 967–974.

Michnicka, M. J., Harper, J. W., & King, G. C. (1993) *Biochemistry 32*, 395–400.

Milligan, J. F., Grobe, D. R., Witherell, G. W., & Uhlenbeck, O. C. (1987) *Nucleic Acids Res. 15*, 8783–8798.

Müller, N., Ernst, R. R., & Wüthrich, K. (1986) *J. Am. Chem. Soc. 108*, 6482–6492.

Nagai, K., Oubridge, C. J., Jessen, T. H., Li, J., & Evans, P. R. (1990) *Nature (London) 348*, 515–520.

Nilges, M., Clore, G. M., & Gronenborn, A. M. (1988a) *FEBS Lett. 239*, 129–136.

Nilges, M., Gronenborn, A. M., Brünger, A. T., & Clore, G. M. (1988b) *Protein Eng. 2*, 27–38.

Plateau, P., & Gueron, M. (1982) *J. Am. Chem. Soc. 104*, 7310–7311.

Puglisi, J. D., & Tinoco, I., Jr. (1989) *Methods Enzymol. 180*, 304–325.

Puglisi, J. D., Wyatt, J. R., & Tinoco, I., Jr. (1990) *Biochemistry 29*, 4215–4266.

Puglisi, J. D., Tan, R., Calnan, B. J., Frankel, A. D., & Williamson, J. R. (1992) *Science 257*, 76–80.

Rould, M. A., Perona, J. J., & Steitz, T. A. (1991) *Nature (London) 352*, 213–218.

Saenger, W. (1983) *Principles of Nucleic Acid Structure*, Springer-Verlag, New York.

Salemink, P. J. M., Swarthof, T., & Hilbers, C. W. (1979) *Biochemistry 18*, 3477–3485.

SantaLucia, J., Jr., & Turner, D. H. (1991) *J. Am. Chem. Soc. 113*, 4313–4322.

SantaLucia, J., Jr., & Turner, D. H. (1993) *Biochemistry* (submitted for publication).

Shaka, A. J., Barker, P. B., & Freeman, R. (1985) *J. Magn. Reson. 64*, 547–552.

Sheline, C. T., Milocco, L. H., & Jones, K. A. (1992) *Genes Dev. 5*, 2508–2520.

Sklenar, V., Miyoshiro, H., Zon, G., & Bax, A. (1986) *FEBS Lett. 208*, 94–98.

Sugimoto, N., Kierzek, R., Freier, S. M., & Turner, D. H. (1986) *Biochemistry 25*, 5755–5759.

Sussman, J. L., Holbrook, S. R., Warrant, R. W., Church, G. M., & Kim, S.-H. (1978) *J. Mol. Biol. 123*, 607–630.

Tidor, B., Irikur, K. K., Brooks, B. R., & Karplus, M. (1983) *J. Biomol. Struct. Dyn. 1*, 231–252.

Varani, G., & Tinoco, I., Jr. (1991a) *J. Am. Chem. Soc. 113*, 9349–9354.

Varani, G., & Tinoco, I., Jr. (1991b) *Q. Rev. Biophys. 24*, 479–532.

Varani, G., Wimberly, B., & Tinoco, I., Jr. (1989) *Biochemistry 28*, 7760–7772.

Varani, G., Cheong, C., & Tinoco, I., Jr. (1991) *Biochemistry 30*, 3280–3289.

White, S. A., Nilges, M., Huang, A., Brünger, A. T., & Moore, P. B. (1992) *Biochemistry 31*, 1610–1621.

Wimberly, B., & Tinoco, I., Jr. (1993) *Biochemistry 32*, 1078–1087.

Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, Wiley, New York.