Thermodynamics of Folding of the RNA Pseudoknot of the T4 Gene *32* Autoregulatory Messenger RNA†

Huawei Qiu,‡,§ Kumar Kaluarachchi,‡,|| Zhihua Du,⊥ David W. Hoffman,⊥ and David P. Giedroc*,‡

Department of Biochemistry and Biophysics, Center for Macromolecular Design, Institute of Biosciences and Technology, Texas A&M University, College Station, Texas 77843-2128, and Department of Chemistry and Biochemistry, University of Texas at Austin, Austin, Texas 78712

Received November 16, 1995; Revised Manuscript Received January 16, 1996⊗

ABSTRACT: Nucleotides U(−67) to C(−40) at the extreme 5′ end of the gene *32* mRNA in bacteriophage T4 have been shown to fold into an RNA pseudoknot proposed to be important for translational autoregulation. The thermal denaturation of three in vitro transcribed RNAs corresponding to the pseudoknot region has been investigated as a function of Mg^{2+} concentration to begin to elucidate the determinants of the structure and stability of this conformation. T4−35 is a 35-nucleotide RNA containing a 5′ G followed by the natural T4 sequence starting with the mature 5′ end of the mRNA, nucleotides A(−71) to C(−38). A 32-nucleotide RNA, termed T4−32, contains the native sequence from U(−67) to C(−40) with 5′-GC and 5′-CA single-stranded regions appended to the 5′ and 3′ ends of the core sequence, respectively. T4−28 contains only the 28 core nucleotides, and the predicted closing U(−67)-A(−52) base pair in stem 1 has been replaced with a phylogenetically allowed G(−67)-C(−52) base pair. Ribonuclease mapping of T4−32 and imino proton NMR experiments of T4−35 show that both sequences adopt a pseudoknotted conformation. At pH 6.9 and 50 mM NaCl, T4−35 and T4−32 RNAs are characterized by a single major melting transition over a wide range of [Mg^{2+}] (0−6 mM). The $\Delta H°$ of unfolding for T4−35 and T4−32 shows a large dependence on Mg^{2+} concentration; the maximum $\Delta H°$ occurs at about 2.0 mM Mg^{2+} with further addition of Mg^{2+} simply increasing the t_m. Investigation of the [Mg^{2+}] dependence of the t_m suggests that a net of one Mg^{2+} ion is released upon denaturation of T4−35 and T4−32 RNAs. Over the entire [Mg^{2+}] range, the $\Delta G°$ (37 °C) for the folding of T4−35 is consistently 1−1.5 kcal mol^{-1} more negative than T4−32 due to a higher stabilization enthalpy for the natural sequence molecule. In contrast to this behavior, T4−28 gives consistently higher t_m's but less negative enthalpies and is destabilized (at 37 °C) by about 0.5−1.5 kcal mol^{-1} relative to T4−32 and by about 2−3 kcal mol^{-1} relative to T4−35, depending upon cation concentration. ^1H NMR experiments suggest that, even in the presence of 4.0 mM Mg^{2+}, T4−28 RNA does not adopt a stable pseudoknotted conformation. These data show that the stability of the pseudoknot in the gene *32* mRNA encoded by the 28-nucleotide core sequence is significantly influenced by the number and nature of the immediately adjacent "single-stranded" 5′ and/or 3′ nucleotides appended to the core structure. These findings are discussed within the context of the structural model for the evolutionarily related phage T2 and T6 gene *32* mRNA pseudoknots presented in the following paper [Du, Z., Giedroc, D. P., & Hoffman, D. W. (1996) *Biochemistry 35*, 4187−4198].

Gene 32 protein (gp32)[1] is the single-strand binding protein (SSB)[1] encoded by bacteriophage T4. It is the prototype member of a large class of SSBs which function as replication and recombination accessory proteins. They are required in stoichiometric amounts and bind cooperatively and largely without sequence specificity to single-stranded nucleic acid sequences (Karpel, 1990). Although single-stranded deoxyribonucleic acids are bound with higher affinity than are analogous RNA molecules (Newport et al., 1981), T4 messenger RNAs would be bound relatively indiscriminately by gp32 and their synthesis repressed if intracellular concentrations of gp32 or other SSB were to reach even modestly high levels (von Hippel et al., 1982).

To overcome this potentially deleterious situation, bacteriophage T4 has evolved a mechanism whereby the intracellular concentration of gp32 is tightly regulated in an autogenous fashion, i.e., gp32 is a translational repressor of its own biosynthesis (Gold et al., 1976; Krisch & Allet, 1982; McPheeters et al., 1988). Gp32 efficiently represses the translation of a reporter gene fused to the 5′ noncoding region of the gene *32* mRNA (McPheeters et al., 1988). A major functional determinant was found to lie at the extreme 5′ end of the gene *32* mRNA leader sequence, which based on phylogenetic data from the related T2 and T6 phages was proposed to fold into an RNA tertiary structural motif termed

† This work was supported by NIH Grant GM42569 (to D.P.G.), Robert A. Welch Foundation Grant A-1295 (to D.P.G.), and NSF Grant MCB-9406065 (to D.W.H.). The NMR instrumentation at the University of Texas was supported by NSF Grant STI-9413770 (to D.W.H.).

* Corresponding author. Phone: 409-845-4231. FAX: 409-862-4718. e-mail: giedroc@bioch.tamu.edu.

‡ Texas A&M University.

§ Present address: Hematology/Oncology Division, Brigham and Women's Hospital, Harvard Medical School, Boston, MA 02115.

|| Present address: Department of Biochemistry and Cell Biology, Rice University, Houston, TX 77251-1892.

⊥ University of Texas at Austin.

⊗ Abstract published in *Advance ACS Abstracts,* March 15, 1996.

[1] Abbreviations: gp32, gene 32 protein; SSB, single-strand binding protein.

0006-2960/96/0435-4176$12.00/0

a pseudoknot (McPheeters et al., 1988). This region of the gene *32* mRNA is required for the extraordinary stability in T4-infected cells (Gorski et al., 1985) and for efficient RNase E-mediated processing which generates the mature 5′ end of the gene *32* mRNA (Carpousis et al., 1989). A model for translational control has been proposed in which the gp32 monomer binds directly to the pseudoknot and, due to cooperative interactions, nucleates cooperative filling of an immediately adjacent unstructured 55-nucleotide AU-rich region extending through the ribosome binding site and the initiation codon of gene *32*. This results in specific repression of the biosynthesis of gene *32* protein (McPheeters et al., 1988; Shamoo et al., 1991). Recent in vitro and in vivo studies in the phage T4 system support the specific involvement of an RNA pseudoknot in translational control (Shamoo et al., 1993) and in RNase E-dependent recognition and processing of the gene *32* containing polycistronic mRNA transcript (Ehretsmann et al., 1992). The strict conservation in the related bacteriophages T2 and T6 of a nucleotide sequence capable of forming an RNA pseudoknot in a genomic region of high sequence diversity, as well as conservation of the structure of the gene *32* 5′ leader in other T-even phages (M1, K3, Ac3, Ox2) lends further support for the biological importance of this RNA pseudoknot (Loayza et al., 1991).

An RNA pseudoknot is a tertiary structural RNA folding motif originally described for the terminal 3′ noncoding regions of the certain plant viral genomic RNAs [for a review, see Pleij and Bosch (1989)]. Of the many folding topologies which are possible for an RNA pseudoknot, the **h**airpin-type (**H**-type) is by far the most common. The H-type pseudoknot fold is an RNA hairpin of S2 base pairs in which single-stranded nucleotides within the loop and continuous with one of the two strands of the helix fold back and form S1 Watson−Crick base pairs with a nearby or far-removed single-stranded region. This creates two helical stems of S1 and S2 base pairs which may stack coaxially to form a pseudocontinuous RNA helical structure of S1 + S2 base pairs. Two inequivalent single-stranded loops then cross the major (of L1 nucleotides) and minor (of L2 nucleotides) grooves of one face of the A-form RNA helix.

Ribonuclease mapping and chemical modification experiments have demonstrated that simple (25−30 nucleotide) designed RNA molecules can adopt a pseudoknotted RNA structure (Wyatt et al., 1990; Puglisi et al., 1988, 1990). The thermodynamic stabilities of a number of designed RNA pseudoknots have been measured and generally reveal that, at equilibrium and 37 °C in the presence of 5 mM Mg^{2+} ion concentration, the pseudoknot is only marginally more or even slightly less stable than its component S1 or S2 hairpin structures (Puglisi et al., 1988; Wyatt et al., 1990). In a systematic study of a set of pseudoknot RNAs which contained common three base pair stem 1 and five base pair stem 2 helical regions, Wyatt et al. (1990) showed that the stabilities of these short pseudoknotted RNAs and propensity to adopt the pseudoknotted conformation appeared to depend strongly on loop size and sequence and Mg^{2+} concentration. The solution structure of one such designed 26-nucleotide RNA obtained from two-dimensional ^1H NMR studies is consistent with the fundamental principles of the H-type pseudoknotted structure (Puglisi et al., 1990).

A 34-nucleotide RNA similar in sequence to the translational frame-shifting site at the *gag-pol* mRNA junction of mouse mammary tumor virus has been shown to form a RNA pseudoknot (Chen et al., 1995; Shen & Tinoco, 1995). However, in contrast to the designed H-type RNA pseudoknots investigated previously (Wyatt et al., 1990; Puglisi et al., 1990), the two helical stems appeared to form a hinged rather than a coaxially stacked helix−helix junction. The pseudoknot within the T4 gene *32* operator and in the related phages T2 and T6 is an excellent model system to obtain structural and thermodynamic data on a fairly complex, natural sequence RNA. It is among the smallest pseudoknots described to date; furthermore, its structure and stability are not likely to be influenced by other long-range tertiary interactions from other regions within the gene *32* mRNA (McPheeters et al., 1988). The T4 sequence is projected to contain a stem 1 of four base pairs, a stem 2 of seven base pairs, a loop 1 containing a single adenylate residue, and a loop 2 of five nucleotides, within 28 contiguous nucleotides corresponding to nucleotides −67 to −40 of the gene *32* mRNA. In this paper, we present the results of RNase structure probing, imino proton NMR, and thermal denaturation experiments using optical absorbance methods to obtain thermodynamic parameters for the folding of three synthetic RNAs based on the sequences from the T4 gene *32* operator. Our thermodynamic findings are discussed within the context of the available structural and thermodynamic data on pseudoknotted RNA molecules (Puglisi et al., 1988; Wyatt et al., 1990; Shen & Tinoco, 1995), in particular, the structural model of the T2/T6 gene *32* mRNA pseudoknot found in the following paper (Du et al., 1996).

MATERIALS AND METHODS

Materials. All buffers were made in distilled and deionized water and treated with the RNase inhibitor diethyl pyrocarbonate (DEPC) before sterilization by autoclaving. DEPC and buffer salts were obtained from Sigma. RNases used for RNA sequencing were obtained in kit form from United States Biochemicals. For structure probing reactions, RNase T1 was obtained from Sigma and RNase V1 was obtained from Pharmacia. All other molecular biologicals were obtained from either New England Biolabs (Boston, MA), Boehringer-Mannheim, or Promega-Fisher Scientific.

Proteins. T7 RNA polymerase was purified from the overexpression plasmid pAR1219 (Studier et al., 1990) transformed into *Escherichia coli* BL21. Both materials were kindly provided by F. Studier (Brookhaven National Laboratory).

RNAs. Preparation of RNA by in Vitro Transcription and Chemical Synthesis. For T4−32, a synthetic plasmid-based DNA template containing a bacteriophage T7 promoter was prepared and inserted into *Eco*RI/*Bam*HI site of the plasmid pUC18 to create pUCPK (Reyes & Abelson,1989; Milligan & Uhlenbeck, 1989). Cleavage of pUCPK with *Bst*NI produced a template DNA suitable for run-off transcription by T7 RNA polymerase according to standard procedures. The crude RNA product was purified either by preparative denaturing polyacrylamide gel electrophoresis or by anion-exchange (Nucleogen DEAE 500−7) HPLC employing a Waters 600E HPLC system developed with a salt gradient from 0 to 1 M KCl over 35 min or from 0.25 to 1 M KCl over 20 min in denaturing conditions (5 M urea in 20 mM potassium phosphate, pH 7.0) at 60 °C. The fractions containing the RNA were collected and loaded onto a C18

reverse-phase Maxi-Clean cartridge (Alltech Associates, Inc., Deerfield, IL), eluted by 50% methanol and dried under vacuum. This RNA was homogeneous by denaturing polyacrylamide gel electrophoresis. T4−28 RNA was purified from in vitro transcription of a fully duplex synthetic DNA template prepared by annealing the two complementary 45-nucleotide oligodeoxyribonucleotides of the appropriate sequence. T4−35 RNA was prepared in exactly the same way as T4−28, except that a partial duplex template (Milligan & Uhlenbeck, 1989) was used for in vitro transcription. For some studies, RNA molecules corresponding to the 3′ hairpin fragment of T4−32 were chemically synthesized by the Gene Technologies Laboratory at Texas A&M University, desalted on a Sephadex G-25 column, and purified using HPLC in a similar way. The RNA was sequenced using the RNase sequencing kits as described below. The 3′ hairpin RNA is 26 nucleotides and has the sequence 5′-OH-AGCUAUGAGGUAAAGUGUCAUAGC-CA-3′.

Methods. RNase Sequencing and Structure-Probing Experiments. RNA molecules were 5′- or 3′-^{32}P-labeled using standard methodologies and gel purified. In vitro synthesized RNA transcripts were dephosphorylated with calf intestinal alkaline phosphatase and rephosphorylated with [γ-^{32}P]ATP and polynucleotide kinase to effect 5′-^{32}P-labeling. Enzymatic RNA sequencing was carried out following the protocol provided by the RNA sequencing kit manufacturer (USB). For structure-specific RNase digestion reactions, the standard structure-probing buffer (SSPB, 10 mM Tris-HCl, pH 7.0, 10 mM $MgCl_2$, and 100 mM KCl) was used (Knapp, 1989). RNases were diluted into SSPB buffer and the optimal concentration for each enzyme was obtained by analytical digestion and analysis of products by electrophoresis. For a typical reaction, 5 μL of 5′-radiolabeled and heat-denatured/refolded RNA (30 000−50 000 cpm) was mixed with 1.5 μL of tRNA carrier (2 mg/mL) and 8.5 μL of SSPB buffer, incubated on ice for 5 min before the addition of the appropriate amount of RNase. Aliquots (one-third of the total volume) were taken out at the indicated times after addition of RNase at 4 °C, mixed with 4 μL of sequencing stop solution, and stored at −20 °C until electrophoresis on a 20% denaturing polyacrylamide gel.

NMR Spectroscopy. NMR spectra were collected at 500 MHz on a two-channel Bruker AMX spectrometer. Proton spectra were collected in 90% H_2O/10% D_2O on samples containing approximately 1 (T4−35) or 16 (T2/T6 RNA) mg of RNA dissolved in 550 μL of 10 mM Na^+/K^+ phosphate buffer, pH 6.0, using a jump and return pulse sequence (Gueron et al., 1991). For each spectrum of T4−35, 1024 scans of 32 768 points were acquired with a sweep width of 12 195 Hz. For the spectrum of the T2/T6 RNA, 256 scans of 32 768 points were acquired with a sweep width of 12 195 Hz.

Thermal Denaturation Experiments. Metal-free RNAs were prepared for melting experiments by subjecting concentrated solutions of RNA to exhaustive dialysis (Spectrapore 3 tubing, 3500 or 12 000 molecular weight cutoff) against 10 mM sodium phosphate, pH 6.9, 50 mM NaCl, and 10 mM EDTA, followed by three to five changes of the same buffer lacking the EDTA (Puglisi & Tinoco, 1989). Samples (350−400 μL) were prepared at the desired RNA and divalent cation concentration using the final dialysate as RNA diluent. In some cases, the RNA samples were then heated to 90 °C for 2 min, slow-cooled at room temperature over the course of 1 h, degassed under vacuum in a desiccator, and then left on ice for 5 min. Reference samples were prepared in exactly the same way with the final dialysate. $MgCl_2$ was added to the indicated total concentration from a stock $MgCl_2$ solution (Johnson-Mathey puratronic grade) whose concentration was determined by atomic absorption spectroscopy on a Perkin-Elmer 2380 AAS operating in the flame mode. All RNA melts were carried out on a Cary 1 scanning spectrophotometer operating in the double beam mode equipped with a temperature controller. Capped 10 mm path-length masked cells were used with the absorption monitored at 260 nm as a function of temperature. The temperature controller was ramped at a ramp rate of 0.5 °C/min from 5−90 or 5−95 °C, with a data point collected every 0.5 °C increment. Some runs were made at a ramp rate of 0.2 °C/min. Samples were typically prepared on ice and loaded into room temperature cuvettes and the cells allowed to equilibrate at 5.0 °C for 40−60 min. The temperature recording reflects the cell block temperature. All melts were done from low to high temperature with a triplicate ramping program to evaluate the reproducibility of the thermal transitions. A 40−60 min equilibration period was introduced between ramp cycles. Thermodynamic parameters were calculated from the first temperature ramp (Puglisi & Tinoco, 1989).

Data Analysis. Text files of the melting data digitized at 0.5 or 0.2 °C increments were imported into Kaleidagraph (Synergy Software, Reading, PA) and the data smoothed (smoothing window = 8 or 3.5 °C). This value was chosen to optimize the signal-to-noise level in the derivative data. Application of smoothing windows slightly larger or smaller (6−10) did not visually greatly affect the data, suggesting that broad transitions were not selectively missed in the reduction. All melts were subjected to two methods of analysis using the program MeltFit (kindly provided by D. E. Draper, Johns Hopkins University) as described (Laing & Draper, 1994) to obtain thermodynamic parameters. In one method, a sloping baseline fit, the smoothed absorbance data were fit directly to the equation:

$$A_{260} = (m_L T + b_L)/(1 + K) + (m_H T + b_H)K/(1 + K)$$

where m_L is the slope of pretransition region baseline, m_H is the slope of posttransition region baseline, and b_L and b_H define the intercepts of the lower and upper baselines at 0 K, respectively, T is the temperature in kelvin, and K is the equilibrium constant for RNA denaturation:

$$K = \exp[\Delta H(1/t_m - 1/T)/R]$$

where ΔH and t_m define the enthalpy and melting temperature of the RNA, respectively. In the second method of analysis, the first derivative of the smoothed absorbance data with respect to temperature was taken ($\partial A/\partial T$) and the data were normalized by dividing the derivative by the absorbance at the specified temperature. Nonlinear least squares fits of the derivative data to one or multiple independent transitions were done as described (Laing & Draper, 1994). Each ith transition is defined by an equilibrium constant K_i, where ΔH_i (enthalpy of the transition), A_i (amplitude), and $t_{m,i}$ (melting temperature) for each ith of n total thermal transitions are optimized according to the equation:

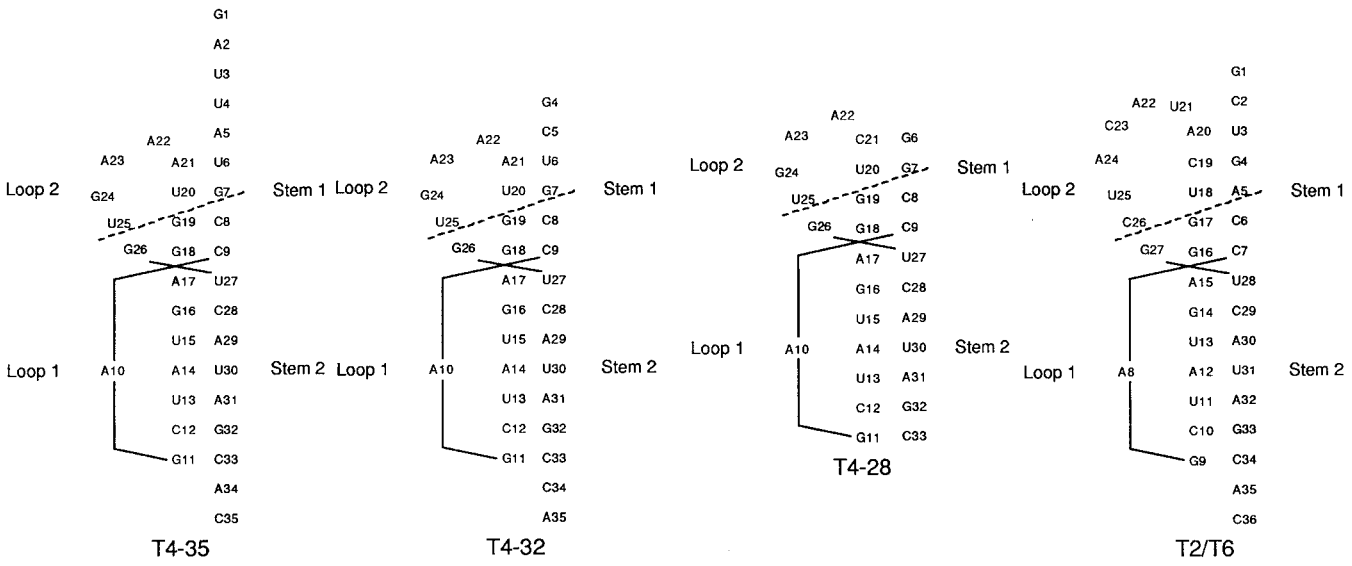

FIGURE 1: Pseudoknot conformational representations of the three T4-derived (T4−35, T4−32, and T4−28) and T2/T6-derived gene *32* translational operator RNAs employed in this study. The nucleotides below the dashed line are identical in both T4- and T2/T6-derived RNAs.

$$\partial A/\partial T = \sum A_i \Delta H_i K_i/[RT^2(1 + K_i)^2]$$

summed from $i = 1$ to n and $K_i = \exp[\Delta H_i(1/t_{m,i} - 1/T)/R]$. Single, double, and triple transition fits were carried out as indicated (Laing et al., 1994). The standard state was taken as 1 M RNA strands, 37 °C, 1 atm. The optimized ΔH was assumed independent of temperature; therefore, $\Delta H = \Delta H°$. $\Delta S°$ was obtained from $\Delta H°/t_m = \Delta S°$, while $\Delta G°$ (37 °C) was from $\Delta G° = \Delta H° - (310.15\ \text{K})\Delta S°$, with $\Delta S°$ also assumed to be temperature-independent. Parameters obtained for both methods of analysis are presented but were generally in good agreement with one another (t_m within 0.5 °C; ΔH within 5−10 kcal mol^{-1}). Thermodynamic parameters resolved from replicate melting experiments using RNA obtained from independent preparations fell within this range (t_m within 0.3 °C; ΔH within 5−10 kcal mol^{-1}).

RESULTS

T4−32 and T4−35 Adopt Pseudoknotted RNA Conformations. The nucleotide sequences (Krisch & Allet, 1982) and folding schemes for the pseudoknotted conformations predicted to be adopted by T4−35, T4−32, and T4−28 RNAs are shown in Figure 1. In the T4−35 RNA, the nucleotides labeled 1−35 correspond to a T7 transcription initiation nucleotide G1 followed by the natural mRNA sequence extending from A(−71) to C(−38) upstream of A(0) in the initiation codon AUG for gene *32*. The most abundant gene *32* mRNA transcripts in T4-infected cells begin at position A(−71) or A(−72) (Mudd et al., 1988). In the T4−32 RNA, the two unpaired nucleotides at the 5′ and 3′ termini are non-T4 sequences and were dictated by the requirements of the plasmid-based in vitro RNA expression system that was employed (Milligan & Uhlenbeck, 1989); the sequence from U6 to C33 corresponds precisely to U(−67) to C(−40) in the native mRNA (McPheeters et al., 1988). In the T4−28 RNA, both 5′ and 3′ "tails" have been removed, requiring introduction of a non-native G6-C21 stem 1 base pair within the 28-nucleotide core sequence. This is expected to be phylogenetically allowed since in both bacteriophages T2 and T6, a G-C base pair is found in an analogous position (McPheeters et al., 1988) although, in the T2/T6 sequence, it is not the terminal base pair of stem 1 (Figure 1).

RNA structure probing with ribonucleases can be used to gain insight into the secondary structure of these RNAs (Knapp, 1989). Although there are clear limitations to the technique, particularly in complex tertiary structure-containing RNAs (Kean & Draper, 1985), duplex structures are usually susceptible to cleavage by the double-strand-specific ribonuclease V1, while single-stranded regions would be readily cleaved by RNase T1, which is also strongly G-specific, and RNase T2, which shows a preference for cleavage to the 3′ side of A nucleotides (Knapp, 1989; McPheeters et al., 1988). The gene *32* mRNA pseudoknot structure predicts a short four base pair stem 1 to form between nucleotides 6−9 and 18−21, with the larger stem 2 formed between nucleotides 11−17 and 27−33. Loop 1 is predicted to be a single A10 while loop 2 would encompass nucleotides A22−G26.

Figure 2 shows a representative structure-mapping experiment carried out with 5′-^{32}P-labeled T4−32 in the presence of 10 mM Mg^{2+} at 4 °C. The most striking feature of this experiment is the relative reactivity of individual nucleotides toward cleavage by RNase T1; G24 in the predicted loop 2 is by far the most reactive nucleotide, with all the other G residues situated in putative stem 1 and stem 2 regions comparably less reactive. Consistent with this, these nucleotides are preferentially cleaved by RNase V1 in T4−32, consistent with formation of the pseudoknot in solution. G7 and C8 are readily cleaved by V1, consistent with the formation of stem 1; however, the situation for nucleotides 18−22 is more complex. While A17, which is predicted to close stem 2, is cleaved preferentially by RNase V1, G18 at the other side of the stem 1−stem 2 junction is not readily cleaved by either T1 or V1. This suggests that this phosphodiester bond is sterically hindered perhaps by the closing of loop 2. G19, on the other hand, is cleaved weakly by both RNases. A21, which is predicted to close stem 1, is cleaved preferentially and weakly by RNase T2 and not at all by RNase V1. This is consistent with the known properties of both RNase T2 and RNase V1. RNase V1 does not cleave well toward helix termini (see the region around

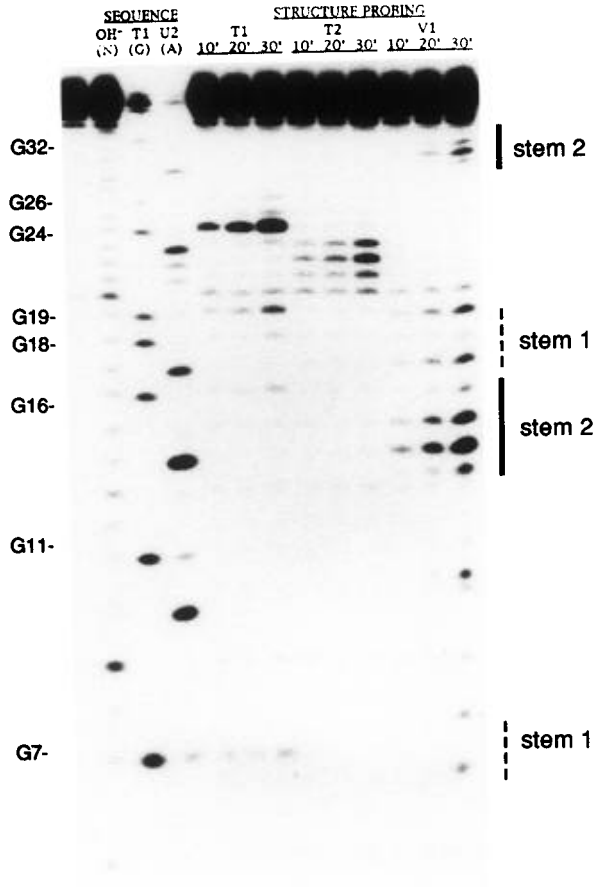

FIGURE 2: Probing of 5′-^{32}P-labeled T4−32 with structure-specific RNases. Refer to Figure 1 for the nucleotide numbering scheme. The leftmost lane is the control RNA lane (no cleavage reagent added). The next three lanes serve as sequencing ladders generated by alkaline hydrolysis (OH^-), RNase T1 (T1), and RNase U2 (U2). The next nine lanes are structure-probing results obtained from the degradation by RNase T1, RNase T2, and RNase V1 as described in Materials and Methods with aliquots taken out following a 10, 20, and 30 min incubation at 4 °C.

G32 in stem 2) (Knapp, 1989). RNase T2, although single-strand specific, is known to cleave in regions of a sharp turn in the polynucleotide backbone adjacent to helical segments (Kean & Draper, 1985). In the pseudoknotted conformation, A21 in T4−32 is likely to be in such a conformation (Du et al., 1996). Clearly, G24, A23, and A22 are among the most reactive toward single-strand-specific ribonucleases, consistent with these nucleotides being unpaired in T4−32. Finally, the unpaired A10 of loop 1 is preferentially cleaved by the double-strand-specific RNase V1. The analogous phosphodiester bond in the pseudoknot formed by the 3′ end of tobacco rattle virus RNA (van Belkum et al., 1987) and the designed pseudoknots (Puglisi et al., 1988) is preferentially cleaved by RNase V1 as well. All of these reactivity characteristics are consistent with those previously published (Shamoo et al., 1993), where S1 nuclease was used as a probe for single-stranded regions of a similar gene *32* mRNA fragment.

We have also structure-probed a 5′-^{32}P-labeled 5′-deletion fragment of T4−32, denoted 3′ hairpin RNA, which extends from nucleotides A10 to A35 (data not shown). In the 3′ hairpin RNA, in contrast to T4−32, G19 and G24 have nearly indistinguishable reactivities toward RNase T1, behavior expected of loop residues. In addition, of the A nucleotides in the A21-A22-A23 sequence, A21 is most

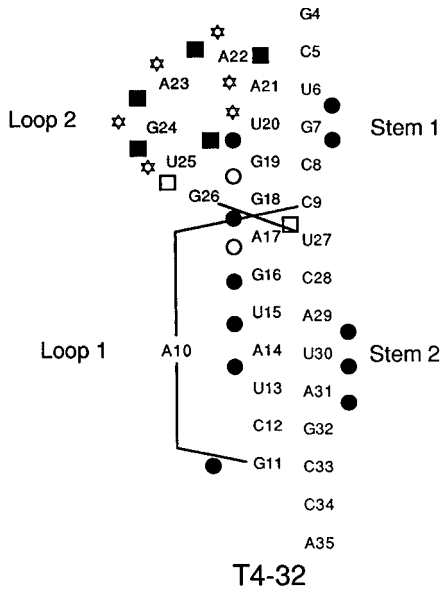

FIGURE 3: Summary of the structure-probing experiments carried out on 5′-^{32}P-labeled T4−32. The phosphodiester bonds susceptible to cleavage by double-strand-specific RNase V1 and single-strand-specific RNases T1 and T2 are given as in the following key: (●) strong V1; (○) weak V1; (■) strong T1/T2; (□) weak T1/T2; (☆) enhanced hydroxyl radical cleavage in the presence of gp32 (data not shown). The stem 1, stem 2, loop 1, and loop 2 regions are indicated.

reactive toward RNase T2 in 3′ hairpin RNA, whereas in the T4−32 RNA, its reactivity is significantly less than that of the A22 and A23 loop 2 residues. G18 and G26 are far less reactive as expected for nucleotides immediately adjacent to a helical segment. Similar behavior is found for these nucleotides in T4−32. These data taken together strongly suggest that a pseudoknotted structure is formed in the T4−32 RNA. Furthermore, the partially folded stem 2 hairpin structure cannot be present to a significant degree (≤10%) in T4−32 at equilibrium under the conditions of 10 mM Mg^{2+} since the reactivity profile (Figure 1) is clearly not a superposition of the profiles characteristic of the pseudoknot and 3′ hairpin conformation, respectively. A summary of the structure-probing experiments carried out with the T4−32 RNA is compiled in Figure 3.

Imino Proton NMR Studies of T4−35. T4−35 RNA was subjected to ^{1}H NMR spectroscopy in 90% H_2O to observe the exchangeable imino protons associated with stable base pairs and some loop residues. If all imino protons exchange sufficiently slowly with solvent, the pseudoknotted conformation of T4−35 should contain 12 imino protons associated with the 11 base pairs of stems 1 and 2 and as many as three additional imino protons associated with loop 2. If only the large stem 2 hairpin forms, a maximum of seven imino protons would be observed.

The imino region of the T4−35 RNA in 10 mM sodium phosphate, pH 6.0, 10 °C is shown in Figure 4A compared to the same region for T2/T6 pseudoknot RNA [see accompanying paper by Du et al. (1996)]. Figure 1 shows that both molecules are identical to one another in stem 2 and in the stem 2−stem 1 junction region and, therefore, would be expected to give a subset of identical resonances in this region. Figure 4A shows that this is the case. T4−35 RNA imino proton assignments derived from extensive residue-specific assignments for T2/T6 pseudoknot are shown (Du et al., 1996). The high correspondence of resonances,

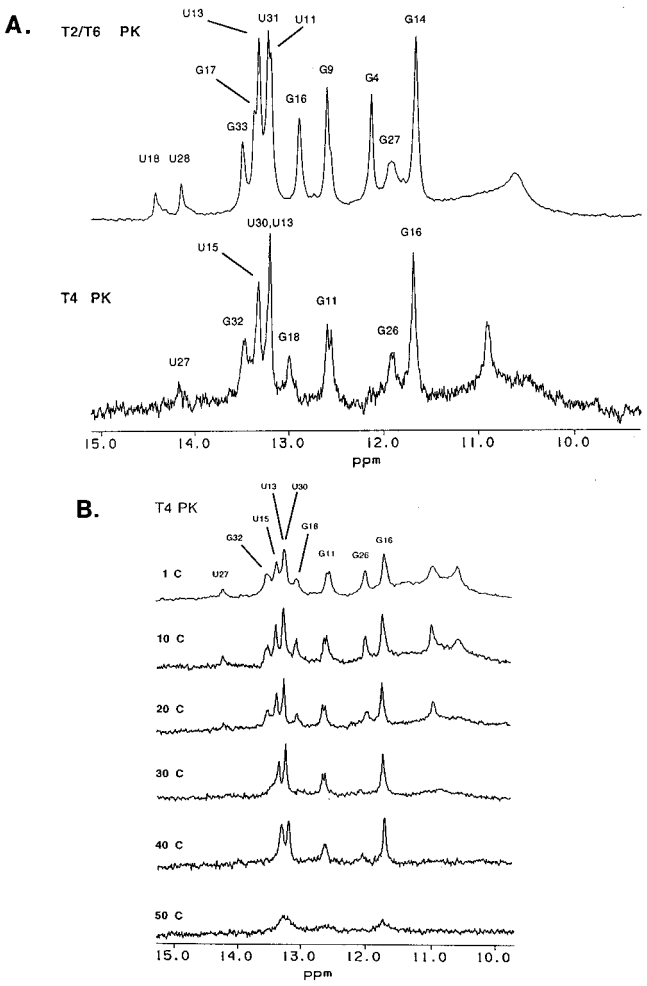

FIGURE 4: (A) Imino proton NMR spectral region of T2/T6 pseudoknot RNA compared with T4−35 RNA. Each spectrum was acquired at 10 °C, 10 mM phosphate, pH 6.0. The resonance assignments shown for T2/T6 RNA are detailed in the following paper (Du et al., 1996). The assignments indicated for T4−35 are preliminary and inferred from the T2/T6 RNA assignments based on chemical shift identity from identical regions of both molecules (see Figure 1). (B) Temperature dependence of the imino proton resonances of T4−35.

in particular U27 which defines the stem 1−stem 2 junction, coupled with the observation of at least 11 imino proton resonances, shows that the pseudoknot conformation forms in the T4−35 RNA.

It is generally agreed that a contributing factor to enhanced rates of imino proton exchange is base pair opening. As the temperature is raised, imino proton resonances broaden as a result of enhanced exchange rates with solvent or via other mechanisms. The temperature dependence of the imino region of T4−35 is shown in Figure 4B. As the temperature is increased, the imino protons disappear in roughly two distinct groups. Clearly, the imino protons associated with the stem 1−stem 2 junction, G26 in loop 2, and base pairs G18-C9 and A17-U27 exchange more rapidly at 1 °C and give rise to weak or broad resonances at low temperature; these protons are lost first as the temperature is increased. At 40 °C, only the central three base pairs of stem 2 as well as the imino proton associated with terminal stem 2 G11-C33 base pair remain, which then broaden at higher temperatures. Analogous findings characterize the T2/T6 pseudoknot (Du et al., 1996).

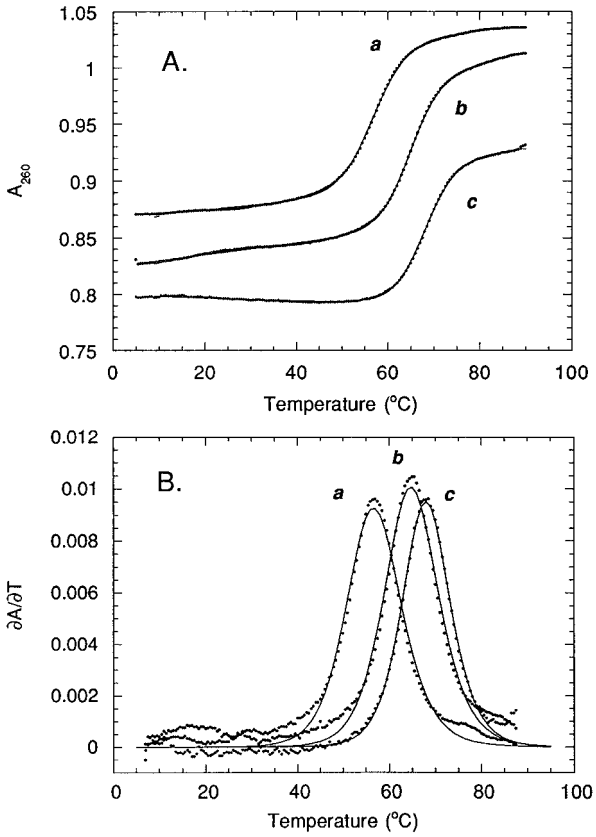

FIGURE 5: (A) Representative thermal denaturation curves for T4−32 RNA at pH 6.9, 50 mM NaCl in the presence of 0 mM Mg^{2+} (no added cation) (curve a), 1.0 mM Mg^{2+} (curve b), and 2.0 mM Mg^{2+} (curve c). Smoothed A_{260} vs temperature data are given by the symbols. The continuous line through the data is the nonlinear least squares optimization of ΔH and t_m to a single two-state transition using the sloping baseline method (see Materials and Methods). The temperature range included in each analysis and optimized values for ΔH and t_m are shown in Table 1. (B) First derivative plots of the same data in panel A, superimposed with a best-fit function for a single transition. The analyzed temperature range and optimized values for ΔH and t_m are given in Table 1.

Thermal Denaturation Studies of T4−35 and T4−32 RNAs. We next used UV absorption spectroscopy to monitor the unfolding of T4−35 and T4−32 RNAs as a function of RNA strand concentration and solution variables. All experiments were carried out in 10 mM sodium phosphate and 50 mM NaCl, pH 6.90, with only the di- or multivalent cation concentration different between samples. Initial melting experiments carried out in the presence of 1.0 mM $MgCl_2$ were done to examine the [T4−32] dependence of the denaturation. As expected for a unimolecular reaction, the t_m was found to be independent of [T4−32] over a 20-fold range extending from 0.35 to 8 μM RNA strand (data not shown). Thermodynamic parameters were obtained over the RNA concentration range of 1−3 μM.

Representative thermal melts obtained with T4−32 and T4−35 RNAs are shown in Figures 5 and 6, respectively, as a function of $MgCl_2$ concentration. In panel A of each figure, the smoothed A_{260} vs temperature profile is shown, while panel B gives the absorbance derivative ($\partial A/\partial T$) plots of the same sets of data. In each panel, the continuous line drawn through the data is a nonlinear least squares fit of the data to the appropriate function (see Materials and Methods) to obtain the optimized parameters summarized in Tables 1 and 2. A_{260} vs temperature plots were directly fit to a unimolecular two-state transition using a sloping baseline

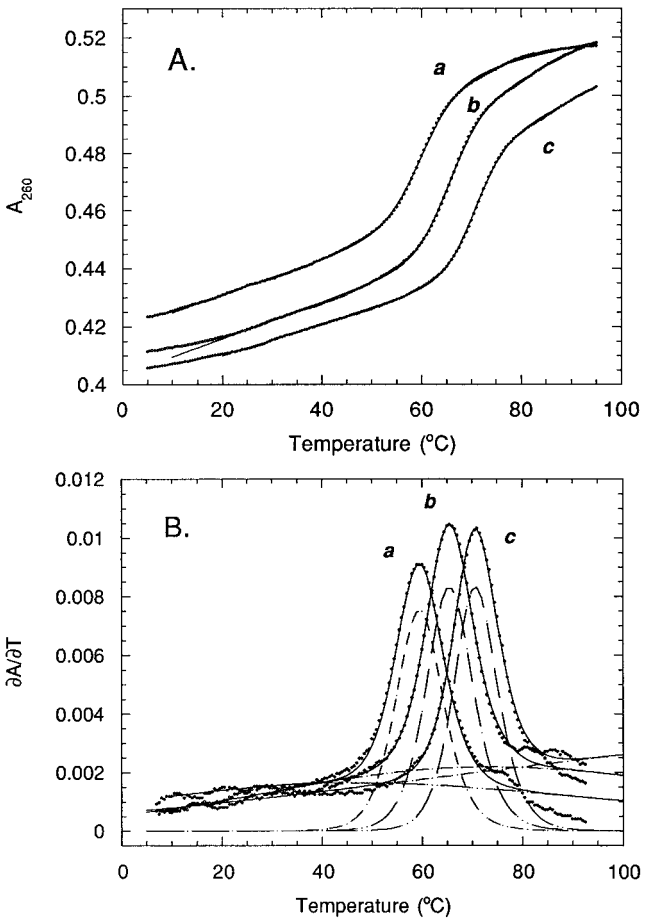

FIGURE 6: (A) Representative thermal denaturation curves for T4−35 RNA at pH 6.9, 50 mM NaCl in the presence of 0 mM Mg^{2+} (no added cation) (curve a), 0.5 mM Mg^{2+} (curve b), and 2.0 mM Mg^{2+} (curve c). As in Figure 5, the smoothed A_{260} vs temperature data are given by the symbols with the solid curve defined by a nonlinear least squares optimization of ΔH and t_m to a single two-state transition using the sloping baseline method (see Materials and Methods). The temperature range, ΔH, and t_m are given in Table 2. (B) Temperature-normalized (T = 20 °C) first derviative plots of the same data in panel A, superimposed with a best-fit function described by two independent transitions. The extremely broad transitions required to fit the non-zero $\partial A/\partial T$ at temperatures below the major transition are not at all well-defined by the data; however, the parameters obtained for the major transition are insensitive to either the enthalpy or t_m defined by this transition. The temperature range included in each analysis and optimized ΔH and t_m values are compiled in Table 2.

method, while the absorbance derivative data were fit assuming one or two independent melting transitions with the parameters ΔH and t_m of each transition having been optimized. If a single-transition fit provides an acceptable fit to the derivative data and resolved parameters are in good agreement (within 10%) with those obtained from the sloping baseline analysis, this is behavior consistent with a two-state unfolding model.

T4−32 and T4−35 RNAs will be discussed in turn. For the T4−32 RNA (Figure 5), these and all other melts not shown appear reasonably well described by a single transition. We could obtain no consistent evidence for the presence of an additional transition over the entire temperature range from 7 to 85 °C under any conditions, even in the absence of divalent cation and when the scan rate was reduced. Standard state thermodynamic parameters have been calculated for the primary melting transition in T4−32 assuming a two-state transition and are compiled in Table 1.[2]

For T4−35 RNA, the melts in general appeared to be more complex than those obtained for T4−32 RNA. Although the sloping baseline method fits always gave an excellent fit to the data (Figure 6A), absorbance derivative plots when fit with a single transition gave poor quality fits and too little enthalpy change. This was due to the consistently non-zero value for $\partial A/\partial T$ at low temperatures. This has also been observed for other complex RNAs (Laing & Draper, 1994). For example, at Mg^{2+} concentrations less than 1.0 mM, the absorbance derivative at low temperatures is small but nearly constant and not obviously a transition (curves a and b in Figure 6B). This probably derives from the rather steep slope in the pretransition baseline region since when this region was fitted as a separate independent melting transition, the t_m and enthalpy change associated with the major transition were found to be in good agreement with parameters obtained with the sloping baseline method of analysis (see Figure 6B, no Mg^{2+}, curve a).[3] Since the enthalpy change obtained for the major transition was largely insensitive to that obtained for the broad transition, it was not considered further. Standard state thermodynamic parameters have been calculated for the major melting transition in T4−35 RNA assuming a two-state transition and are compiled in Table 2.

Inspection of Tables 1 and 2 reveals that under all solution conditions the 5′-extended T4−35 RNA molecule consistently melts with a t_m 2.5−3.0 °C larger than T4−32. Furthermore, the melting of T4−35 returns a consistently larger negative enthalpy change (by about 5−10 kcal mol^{-1}) and entropy change for the folding reaction, relative to T4−32 RNA. This results in a $\Delta G°$ (37 °C) which is consistently 1−1.5 kcal mol^{-1} more negative for T4−35 relative to T4−32, depending upon conditions. Thus, the nature and/or number of nucleotides at the 5′ and/or 3′ end of the core 28-nucleotide pseudoknot sequence significantly influences the overall thermodynamic stability of the folded conformation. Although the structural origin of the higher stability of T4−35 is not yet known, the magnitude of the difference is not at all unreasonable since nearest neighbor thermodynamics predicts that, by simply replacing the native A34 in T4−35 with C34 in T4−32, the enthalpy of denaturation may be reduced by as much as 5 kcal mol^{-1} in 1 M Na^+ (Turner et al., 1988).

As expected from previous experiments on a number of designed RNA pseudoknots (Wyatt et al., 1990), both T4−35 and T4−32 RNAs are strongly stabilized by Mg^{2+} binding. For example, the t_m for T4−32 increases from 57 °C in the absence of cation to 73 °C in the presence of 6

[2] The thermodynamic parameters which describe the folding of T4−32 and T4−35 RNAs differ significantly from those obtained for a 25 nucleotide 3′ hairpin RNA. For example, in the presence of 0, 1.0, and 2.0 mM Mg^{2+}, the t_m and enthalpy of unfolding of this RNA are 61.5 °C, 50 kcal mol^{-1} (0 mM Mg^{2+}), 71.8 °C, 43 kcal mol^{-1} (1.0 mM Mg^{2+}), and 71.9 °C, 62 kcal mol^{-1} (2.0 mM Mg^{2+}).

[3] Although each of the putative broad transitions was characterized by a very low enthalpy, the returned amplitudes and t_m's were consistently anomalously high as to question their significance given the expected correlation between enthalpy and hyperchromicity (Tinoco, 1960). In addition, fitting the derivative data to three independent transitions always gave a slight improvement in the goodness-of-fit parameter but returned too much total enthalpy (110−120 kcal mol^{-1}) to be reasonable (data not shown).

Table 1: Summary of Thermodynamic Parameters for the Folding of T4−32 RNA[a]

$MgCl_2$	temp range (°C)[b]	$\Delta H°$ (kcal mol^{-1})	t_m (°C)	$\Delta S°$ (eu)[c]	$\Delta G°$ (37 °C)[c] (kcal mol^{-1})
0	7−85	−55	56.8	−167	−3.3
	7−90	−58	57.0	−176	−3.5
0.2	7−85	−60	59.5	−180	−4.1
	7−90	−61	59.3	−184	−4.1
0.5	7−85	−61	63.0	−181	−4.7
	7−90/30−90[d]	−64 (±6)	62.7	−190	−4.9
1.0	7−85	−58	64.9	−172	−4.8
	8−90	−67	64.9	−198	−5.5
2.0	7−85	−67	68.1	−196	−6.1
	16−90	−69	67.9	−201	−6.2
3.0	7−85	−67	70.3	−195	−6.5
	23−90	−71	70.3	−205	−7.1
6.0	7−85	−66	72.5	−191	−6.8
	23−90	−71	73.1	−204	−7.4
2.0 mM spermidine	7−85	−73	73.6	−211	−7.5
	7−90	−68	73.7	−197	−7.2

[a] Conditions: 10 mM sodium phosphate, 50 mM NaCl, pH 6.9, and the indicated cation concentration. In all cases, the absorbance at 260 nm was recorded as a function of temperature from 5 to 90 °C at a ramp rate of 0.5 °C/min. The first entry describes fitted parameters (ΔH and t_m) derived from nonlinear least squares analysis of $\partial A/\partial T$ vs T data like that shown in Figure 5B. Single transition fits are reported in all cases. The second entry represents an analysis of the A_{260} vs temperature data using the sloping baseline method (see Figure 5A and Materials and Methods). The range in ΔH and t_m from multiple experiments is ±5 kcal mol^{-1} and ±0.3 °C, respectively, unless otherwise indicated. [b] Temperature range included in the analysis. [c] $\Delta S°$ was obtained from $\Delta H°/t_m = \Delta S°$, while $\Delta G°$ (37 °C) was from $\Delta G° = \Delta H° - (310.15\ \text{K})\Delta S°$, with $\Delta S°$ and $\Delta H°$ assumed to be temperature-independent. The uncertainty in $\Delta G°$ (37 °C) is ±0.4 kcal mol^{-1}. [d] The range of parameters obtained from the two temperature ranges analyzed is given.

Table 2: Summary of Thermodynamic Parameters for the Folding of T4−35 RNA[a]

$MgCl_2$	temp range[b]	$\Delta H°$ (kcal mol^{-1})	t_m (°C)	$\Delta S°$ (eu)[c]	$\Delta G°$ (37 °C)[c] (kcal mol^{-1})
0	7−92.5	−70	59.6	−210	−4.8
	20−95	−58	60.4	−173	−4.1
0.2	7−92.5	−69	62.8	−205	−5.3
	20−95	−58	63.2	−172	−4.5
0.5	7−92.5	−73	65.6	−216	−6.1
	20−95	−67	65.5	−198	−5.6
1.0	7−92.5	−75	68.1	−220	−6.8
	20−95	−73	68.2	−214	−6.7
2.0	7−92.5/7−82[d]	−78	70.9	−236	−7.9
	20−95	−77	70.5	−224	−7.5
3.0	7−92.5/7−85[d]	−78	72.4	−229	−8.1
	20−95	−79	72.1	−228	−8.0

[a] Conditions: 10 mM sodium phosphate, 50 mM NaCl, pH 6.9, and the indicated cation concentration. In all cases, the absorbance at 260 nm was recorded as a function of temperature from 5 to 95 °C at a ramp rate of 0.5 °C/min. The first entry describes fitted parameters (ΔH and t_m) derived from nonlinear least squares analysis of $\partial A/\partial T$ vs T data like that shown in Figure 6B. These data were fitted to two independent transitions with parameters shown for the major transition only. The second entry represents an analysis of the A_{260} vs temperature data using the sloping baseline method to a single two-state transition (see Figure 6A). [b] Temperature range included in the analysis. [c] $\Delta S°$ and $\Delta G°$ (37 °C) were obtained as described in Table 1. The uncertainty in $\Delta G°$ (37 °C) is ±0.4 kcal mol^{-1}. [d] The parameters recorded span the range obtained upon analysis of the two temperature regions shown.

mM $MgCl_2$. An approximation of the net number of Mg^{2+} ions released (n) as a result of thermal unfolding of T4−32 can be estimated from a change in t_m as a function of Mg^{2+} concentration according to the following equation (Puglisi & Tinoco, 1989):

$$\partial(1/t_m)/\partial \ln [MgCl_2] = -n(R/\Delta H^o)$$

where $\Delta H°$ is the average enthalpy of unfolding and R is the gas constant. A plot of $\partial(1/t_m)/\partial \ln [MgCl_2]$ is shown in Figure 7 for both T4−32 and T4−35 RNAs. For T4−32, the plot is linear over the Mg^{2+} concentration range examined with an $n = 1.07$ (ranging from 0.96 to 1.17 which spans the range in $\Delta H°$ values observed). This suggests that the net binding of a single Mg^{2+} ion to the folded conformation of T4−32 is responsible for the enhanced stabilization by Mg^{2+}. T4−35 exhibits exactly analogous behavior with an $n = 0.99$ (Figure 7), providing further support for structural similarity of these molecules. Spermidine^{3+} can also strongly stabilize T4−32 against thermal denaturation, suggesting that other multivalent cations can substitute for Mg^{2+} in this stabilization. Tables 1 and 2 reveal that the standard enthalpy of formation ($\Delta H°$) of T4−32 is about −58 kcal mol^{-1} in the absence of Mg^{2+} and reaches about −70 kcal mol^{-1} in the presence of 6 mM Mg^{2+} or 2 mM spermidine^{3+}. For T4−35, the enthalpy increases from −60 to near −80 kcal mol^{-1} in the presence of 3 mM Mg^{2+}. Thus, a substantially increased enthalpy of formation underlies the thermodynamic driving force for the greater stability of both T4−35 and T4−32 RNAs in the presence of divalent cations.

T4−28 RNA Lacks the 5′ and 3′ Unpaired Nucleotides and Is Destabilized Relative to T4−32 and T4−35 RNAs. In the mature gene *32* mRNA, the core 28-nucleotide pseudoknot structure is flanked by a short 4-nucleotide unpaired region on the 5′ side and a long AU-rich region on the 3′ side (McPheeters et al., 1988). Since unpaired nucleotides, particularly on the 3′ side of an RNA helical region, can have significant effects on the stability of the

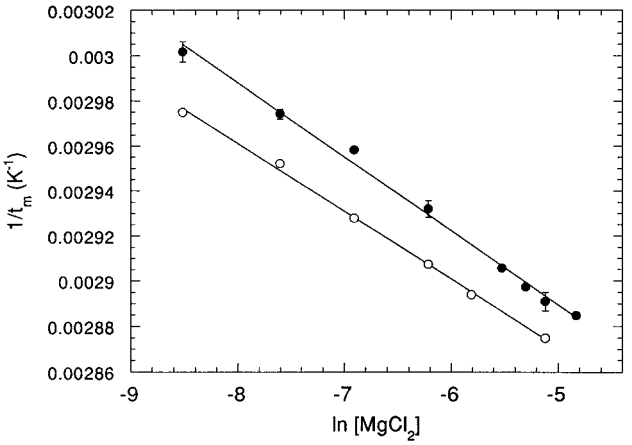

FIGURE 7: The Mg^{2+} dependence of the thermal denaturation of T4−32 (●) and T4−35 (○) RNAs as expressed in a plot of $\partial(1/t_m)/\partial \ln [MgCl_2]$. The solid line represents a linear fit to the equation $\partial(1/t_m)/\partial \ln [MgCl_2] = -n(R/\Delta H°)$. For T4−32, using $\Delta H° = 60$ kcal mol^{-1} gives $n = 1.02$, while for T4−35, using $\Delta H° = 65$ kcal mol^{-1} gives $n = 0.99$.

helix (Sugimoto et al., 1987), we were interested in investigating whether these flanking sequences significantly stabilize or otherwise alter the stability of the gene *32* mRNA pseudoknot.

Representative thermal melting experiments obtained with the T4−28 RNA are shown in Figure 8 at 0, 1.0 mM Mg^{2+}, and 2.0 mM spermidine^{3+}. In contrast to the behavior of T4−35 and T4−32 RNAs, the T4−28 RNA appears to melt in a complex non-two-state fashion at Mg^{2+} concentrations less than about 0.5 mM (0.1, 0.2, 0.35, and 0.5 mM Mg^{2+} were investigated). For example, the derivative plot of the melting of T4−28 in the absence of Mg^{2+} appears to have a shoulder at ≈45 °C, indicative of one or multiple intermediate conformations (Figure 8B). Consistent with this, if the melting curve itself is analyzed applying the two-state assumption as in Figure 8A, an unreasonably low enthalpy (−33 kcal mol^{-1}) is obtained. Attempts to fit the absorbance derivative data obtained at no or low Mg^{2+} concentration with two independent transitions result in a significant improvement in the quality of the fit but return far too much enthalpy and unreasonable relative transition amplitudes than can be rationalized on the basis of the T4−28 base-stacking scheme shown in Figure 1. Although a complete analysis was not undertaken, the shape of individual melting curves did not change when the RNA concentration was varied by a factor of 3, suggesting that the observed melting behavior is reporting on the unfolding of a monomeric species. In contrast to the situation at low Mg^{2+}, at Mg^{2+} concentrations exceeding 1.0 mM or in the presence of 2.0 mM spermidine, a melting transition which is well described as two-state is obtained (Figure 8B). Under these conditions, the t_m of T4−28 is consistently *higher*, but the $\Delta H°$ is consistently *lower* by about 10 kcal mol^{-1} than found for T4−32 and 15−20 kcal mol^{-1} than for T4−35 under the same solution conditions. Standard state thermodynamic parameters are compiled for T4−28 in Table 3 as a function of Mg^{2+} assuming a single two-state transition. T4−28 is characterized by a consistently smaller stabilization free energy by ≈0.8 kcal mol^{-1} relative to T4−32 an ≈1.5 kcal mol^{-1} relative to T4−35.

One- and two-dimensional ^{1}H NMR spectroscopy of the T4−28 RNA was carried out to probe its conformation under

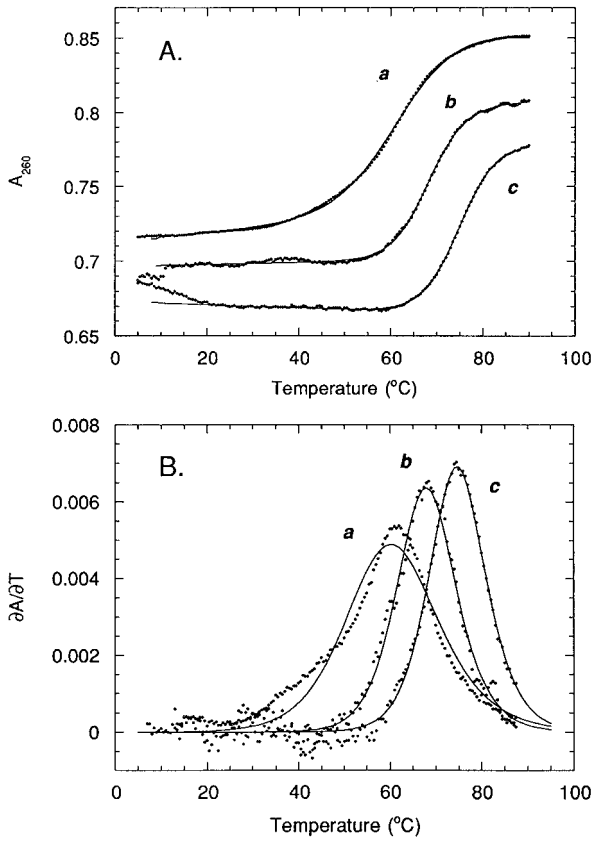

FIGURE 8: (A) Representative thermal denaturation curves for T4−28 RNA at pH 6.9, 50 mM NaCl in the presence of 0 mM Mg^{2+} (no added cation) (curve a), 1.0 mM Mg^{2+} (curve b), and 2.0 mM spermidine^{3+} (curve c). The raw A_{260} vs temperature data are superimposed with a continuous line defined by a nonlinear least squares optimization of ΔH and t_m to a single two-state transition using the sloping baseline method (see Materials and Methods). The temperature range included in each analysis and optimized ΔH and t_m values are summarized in Table 3. (B) First derivative plots of the same data in panel A, superimposed with a best-fit function described by a single independent transition. In curves b and c, $\partial A/\partial T$ values between 7.5 and 12.5 °C and 7.5 and 29 °C, respectively, are not shown and are regarded as noise on the basis of the analysis of replicate experiments. The analyzed temperature range and optimized ΔH and t_m values are compiled in Table 3.

solution conditions similar to the melting experiments. Spectra obtained in 4 mM Mg^{2+} and 25 °C revealed an imino proton spectrum quite distinct from that of T4−35; in addition, 2D NOESY spectra collected in D_2O were not readily interpretable in terms of a single well-defined structure, in contrast to T4−35.[4]

DISCUSSION

In this report, we have compared the structure and thermodynamic stabilities of three in vitro transcribed RNAs based on the nucleotides −71 to −38 in the 5′ noncoding region of the T4 gene *32* mRNA. Previously reported RNase structure mapping experiments have provided evidence consistent with the 28-nucleotide core sequence extending from −67 to −40 adopting an RNA pseudoknot (McPheeters et al., 1988; Shamoo et al., 1993). The T4−35 RNA contains the natural sequence, while the T4−32 and T4−28 RNAs are sequentially shortened versions of the natural sequence (Figure 1). We provide evidence that T4−35 and T4−32

[4] K. Kaluarachchi and D. P. Giedroc, unpublished observations.

Table 3: Summary of Thermodynamic Parameters for the Folding of T4−28 RNA[a]

$MgCl_2$	temp range[b]	$\Delta H°$ (kcal mol^{-1})	t_m (°C)	$\Delta S°$ (eu)[c]	$\Delta G°$ (37 °C)[c] (kcal mol^{-1})
0	7−85	−33	60.8	−100	−2.4
	7−90	−32	62.8	−96	−2.6
0.2	7−85	−34	63.5	−100	−2.8
	7−90	−38	64.5 (±0.6)	−113	−3.0
0.5	7−85	−46	70.2	−135	−4.6
	30−90	−58	69.6	−190	−4.9
1.0	7−85	−54	68.1	−159	−4.9
	13−90	−58	68.4	−170	−5.3
2.0	7−85	−50	70.7	−145	−4.9
	7−70/30−90[d]	−53 (±8)	71.0	−155 (±25)	−5.4 (±0.7)
3.0	7−85	−47	71.5	−137	−4.7
	8−90	−58	71.7	−170	−5.7
4.0	7−85	−56	73.0	−161	−5.8
	16−90	−59	72.7	−171	−6.1
6.0	7−85	−56	74.0	−161	−5.8
	21−90	−63	74.1	−182	−6.7
2.0 mM spermidine	7−85	−59	74.9	−169	−6.4
	21−90	−59	75.2	−171	−6.4

[a] Conditions: 10 mM sodium phosphate, 50 mM NaCl, pH 6.9, and the indicated cation concentration. In all cases, the absorbance at 260 nm was recorded as a function of temperature from 5 to 90 °C at a ramp rate of 0.5 °C/min. The first entry describes fitted parameters (ΔH and t_m) derived from nonlinear least squares analysis of $\partial A/\partial T$ vs T data like that shown in Figure 8B. Single transition fits are reported in all cases. The second entry represents an analysis of the A_{260} vs temperature data using the sloping baseline method (see Figure 8A). [b] Temperature range included in the analysis. [c] $\Delta S°$ and $\Delta G°$ (37 °C) were obtained as described in Table 1. The uncertainty in $\Delta G°$ (37 °C) is ±0.4 kcal mol^{-1} unless otherwise indicated. [d] The range of parameters obtained from the two temperature regions shown is given.

RNAs fold into a pseudoknotted RNA structure as evidenced by ribonuclease structure mapping, imino proton NMR studies, and thermal denaturation experiments.

Mg^{2+} strongly stabilizes the folded form of all three molecules, increasing the t_m by as much as 16 °C over the range of Mg^{2+} tested and increasing the $\Delta H°$ of unfolding by 10−15 kcal mol^{-1} depending upon the molecule. Although striking, this is in general not surprising and cannot be taken as evidence that a complex tertiary structure is formed in these RNAs; simple RNA hairpins are also strongly stabilized by Mg^{2+} binding (Laing et al., 1994). Interestingly, Mg^{2+} does not appear to be *required* for formation of the pseudoknotted tertiary structure in T4−32 or T4−35 RNAs but merely stabilizes the folded conformation. The most compelling evidence in support of this is that the imino proton spectrum of the T4−35 RNA, collected in the absence of added Mg^{2+}, clearly signifies a pseudoknot conformation on the basis of the spectral similarity between T4−35 and the extensively characterized T2/T6 pseudoknot reported in the following paper (Du et al., 1996). Analysis of the Mg^{2+} dependence of the t_m according to a simple nonspecific binding model suggests that the folded form binds a net of one additional Mg^{2+} relative to the unfolded form (Figure 7). The simplest interpretation of these data is that one Mg^{2+} ion binds to a region of high charge density perhaps created by close apposition of multiple phosphodiester bonds, particularly at the loop−stem junctions (Puglisi et al., 1990), thereby stabilizing the pseudoknot structure, perhaps by enhancing base−base interactions. The finding that spermidine^{3+} substitutes effectively for Mg^{2+} suggests that coordination properties of the cation do not strongly effect this stabilization, thereby ruling out a formally site-specific binding model (Laing et al., 1994).

At the highest multivalent cation concentration tested, the observed enthalpies of folding of T4−32 (≈−70 kcal mol^{-1}) and T4−35 (−75 to −80 kcal mol^{-1}) RNAs are both less than that predicted for a continuous 11 base pair duplex formed by the coaxial stacking of stems 1 and 2 (−97 kcal mol^{-1} in 1 M Na^+, pH 7.0) (Turner et al., 1988; He et al., 1991). Several designed pseudoknotted RNAs behave in qualitatively the same manner (Wyatt et al., 1990; Puglisi et al., 1988). One unlikely explanation for this is that the major melting transition which was quantitatively analyzed to obtain the enthalpy of folding contains only contributions from the stem 2 structure. This seems unlikely since all three molecules investigated have the same stem 2 sequence and yet give easily distinguishable enthalpies of denaturation and melting temperatures. Another more likely possibility is that the folding of the pseudoknot, particularly in the stem−loop and stem−stem junction regions, slightly alters favorable short range base−base stacking interactions, thereby diminishing the enthalpic stabilization of the molecule (Wyatt et al., 1990).

The 28-nucleotide T4−28 RNA contains only the core pseudoknot sequence. When the major melting transition of T4−28 is analyzed as a single two-state transition, the structure which is formed, even at high cation concentration, is significantly destabilized relative to T4−32 and more so relative to T4−35. In fact, the enthalpy of denaturation of the major transition never exceeds that predicted for the stem 2 hairpin alone (59 kcal mol^{-1}) if 6 mM Mg^{2+} mimics the conditions of 1 M Na^+ (Turner et al., 1988). These findings coupled with results from ^{1}H NMR spectroscopy of T4−28 RNA[4] suggest that the stem 2 3′ hairpin structure may be the major conformer formed by T4−28, unlike T4−32 or T4−35 RNAs. The destabilization of T4−28 very likely derives from the lack of terminal 5′ and/or 3′ unpaired nucleotides since a comparison of T4−35 vs T4−32 RNAs reveals that the nature and/or number of the 5′ and/or 3′ terminal nucleotides contribute significantly to the stability of the pseudoknot. The solution structure of the T2/T6 mRNA pseudoknot reveals that the 3′ penultimate nucleotide in the single-stranded tail stacks against the terminal base pair of stem 2 and extends the helical sense of the molecule (Du et al., 1996). No such interaction is found at the 5′ end of the molecule. From knowledge of the sequence dependence of the energetics of a 3′ vs 5′ single-stranded nucleotide adjacent to a helical segment in model RNA duplexes, the

3′ A34 stacked against the G11-C33 base pair in the gene *32* mRNA pseudoknot is expected to strongly stabilize stem 2 (Turner et al., 1988). Indeed, we have found that deletion of the 3′ single-stranded nucleotides destabilizes the T2/T6 pseudoknot as evidenced by imino proton exchange data.[5] As the terminal G11-C33 base pair in T4−35 defines the loop 1 A10−stem 2 junction, the additional helical stabilization provided by the 3′ single-stranded stack may play a major role in stabilizing the overall tertiary fold of the molecule. Indeed, the imino proton of G11 (and the analogous nucleotide G9 in T2/T6 RNA) is among the slowest exchanging in both molecules, behavior unexpected for a terminal base pair.

The studies presented here with variant T4 pseudoknot sequences suggest that the introduction of a compensatory base change within defined helical stem regions and/or substitution or deletion of immediately adjacent single-stranded nucleotides may have profound effects on the structure and stability of the pseudoknotted RNA conformation. A systematic investigation of sequence variants within the context of the native operator defined by T4−35 coupled with investigation of the thermodynamic stabilities and solution structures of T4 and T2/T6 gene *32* mRNA operators (Du et al., 1996) is required to fully determine the basis for the lower stability of the T4−28 RNA relative to the natural operator. These studies are currently in progress in our laboratories.

REFERENCES

Carpousis, A. J., Mudd, E. A., & Krisch, H. M. (1989) *J. Mol. Biol. 183*, 165−177.

Chen, X., Chamorro, M., Lee, S. I., Shen., L. X., Hines, J. V., Tinoco, I., Jr., & Varmus, H. E. (1995) *EMBO J. 14*, 842−852.

Du, Z., Giedroc, D. P., & Hoffman, D. W. (1996) *Biochemistry 35*, 4187−4198.

Ehretsmann, C. P., Carpousis, A. J., & Krisch, H. M. (1992) *Genes Dev. 6*, 149−159.

Gold, L., O'Farrell, P. Z., & Russel, M. (1976) *J. Biol. Chem. 251*, 7251−7262.

Gorski, K., Roch, J.-M., Prentki, P., & Kirsch, H. M. (1985) *Cell 43*, 461−469.

Gueron, M., Plateau, P., & Decorps, M. (1991) *Prog. Nucl. Magn. Reson. Spectrosc. 23*, 135−219.

He, L., Kierzek, R., SantaLucia, J., Walter, A. E., & Turner, D. H. (1991) *Biochemistry 30*, 11124−11132.

Karpel, R. L. (1990) in *The Biology of Nonspecific DNA−Protein Interactions* (Revzin, A., Ed.) pp 103−130, CRC Press, Inc., Boca Raton, FL.

Kean, J. M., & Draper, D. E. (1985) *Biochemistry 24*, 5052−5061.

Knapp, G. (1989) *Methods Enzymol. 180*, 192−212.

Krisch, H. M., & Allet, B. (1982) *Proc. Natl. Acad. Sci. U.S.A. 79*, 4937−4941.

Laing, L. G., & Draper, D. E. (1994) *J. Mol. Biol. 237*, 560−576.

Laing, L. G., Gluick, T. C., & Draper, D. E. (1994) *J. Mol. Biol. 237*, 577−587.

Loayza, D., Carpousis, A. J., & Krisch, H. M. (1991) *Mol. Microbiol. 5*, 715−725.

McPheeters, D. S., Stormo, G. D., & Gold, L. (1988) *J. Mol. Biol. 201*, 517−535.

Milligan, J. F., & Uhlenbeck, O. C. (1989) *Methods Enzymol. 180*, 51−62.

Mudd, E. A., Prentki, P., Belin, D., & Krisch, H. M. (1988) *EMBO J. 7*, 3601−3607.

Newport, J. W., Lonberg, N., & Kowalczykowski, S. C. (1981) *J. Mol. Biol. 145*, 105−121.

Pleij, C. W. A., & Bosch, L. (1989) *Methods Enzymol. 180*, 289−303.

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

Puglisi, J. D., Wyatt, J. R., & Tinoco, I., Jr. (1988) *Nature 331*, 283−286.

Puglisi, J. D., Wyatt, J. R., & Tinoco, I., Jr. (1990) *J. Mol. Biol. 214*, 437−453.

Reyes, V. M., & Abelson, J. N. (1989) *Methods Enzymol. 180*, 63−69.

Shamoo, Y., Webster, K. R., Williams, K. R., & Konigsberg, W. H. (1991) *J. Biol. Chem. 266*, 7967−7970.

Shamoo, Y., Tam, A., Konigsberg, W. H., & Williams, K. R. (1993) *J. Mol. Biol. 232*, 89−104.

Shen, L. X., & Tinoco, I., Jr. (1995) *J. Mol. Biol. 247*, 963−978.

Studier, F. W., Rosenberg, A. H., Dunn, J. J., & Dubendorff, J. W. (1990) *Methods Enzymol. 185*, 60−89.

Sugimoto, N., Kierzek, R., & Turner, D. H. (1987) *Biochemistry 26*, 4554−4558.

Tinoco, I., Jr. (1960) *J. Am. Chem. Soc. 82*, 4785−4790.

Turner, D. H., Sugimoto, N., & Freier, S. M. (1988) *Annu. Rev. Biophys. Biophys. Chem. 17*, 167−192.

van Belkum, A., Cornelissen, B., Linthorst, H., Bol, J., Pley, C., & Bosch, L. (1987) *Nucleic Acids Res. 15*, 2837−2850.

von Hippel, P. H., Kowalczykowski, S. C., Lonberg, N., Newport, J. W., & Paul, L. S. (1982) *J. Mol. Biol. 162*, 795−818.

Wyatt, J. R., Puglisi, J. D., & Tinoco, I., Jr. (1990) *J. Mol. Biol. 214*, 455−470.

[5] Z. Du and D. W. Hoffman, unpublished observations.

BI9527348