DNA and RNA Oligomer Sequences from the 3′ Noncoding Region of the Chicken Glutamine Synthetase Gene Form Intramolecular Hairpins†

Peter V. Riccelli,‡ Jovencio Hilario,‡ Frank J. Gallo,‡ Anthony P. Young,§ and Albert S. Benight*,‡

Department of Chemistry, m/c 111, 845 West Taylor Street, Room 4500, University of Illinois, Chicago, Illinois 60607, and Biotechnology Center, The Ohio State University, Rightmire Hall, Columbus, Ohio 43210

Received June 24, 1996; Revised Manuscript Received September 19, 1996⊗

ABSTRACT: The DNA sequence of the chicken glutamine synthetase gene contains an A·T-rich stretch of approximately 1500 base pairs in the 3′ noncoding regions of exon 7 [Pu, H., & Young, A. P. (1989) *Gene 18*, 169−175]. Within this region several palindromic sequences occur that could conceivably form intramolecular structures. One such perfect inverted repeat sequence resides between positions 2605 and 2623. To investigate the hairpin-forming potential for this sequence, optical and calorimetric melting and gel electrophoresis studies have been performed on the following synthetically prepared DNA and RNA oligomer subsequences: DNA, 5′d-T-T-T-T-T-T-A-A-T-A-A-T-T-A-A-A-A-A-A-3′; and RNA, 5′r-U-U-U-U-U-U-A-A-U-A-A-U-U-A-A-A-A-A-A-3′. The DNA strand corresponds to the coding strand sequence while the RNA strand represents the transcribed mRNA. Results of melting analysis of these 19-base, partially self-complementary strands performed in 115 mM Na^+ yielded evaluations of their thermodynamic transition parameters. These values are consistent with the melting of unimolecular structures, presumably hairpins. Thermodynamic parameters evaluated by analysis of the optical melting transitions assuming a two-state model and measured directly by differential scanning calorimetry agreed within experimental error. Therefore, melting behavior of the hairpins is all-or-none like. The DNA hairpin is slightly more stable than the RNA hairpin with melting enthalpy $\Delta H^0 = 41.2 \pm 3.8$ kcal/mol and entropy $\Delta S^0 = 133 \pm 11$ cal/K·mol (eu) compared to $\Delta H^0 = 32.0 \pm 6.0$ kcal/mol and entropy $\Delta S^0 = 105 \pm 20$ eu for the RNA. Gel electrophoretic analysis of these oligomers alone and in various mixtures with their DNA and RNA complementary strands was also performed. Consistent with interpretations of melting results, these experiments revealed both strands alone preferentially form intramolecular hairpin structures. In mixtures in which their complementary strands are in vast molar excess (stoichiometric ratios > 10:1), the intramolecular structures are converted to intermolecular duplexes. For the DNA and RNA strands examined, the conversion is not complete until over a 1000-fold excess of the complementary strand is added. Semiquantitative analysis of gel electrophoretograms enabled evaluations of the relative free energies of the hairpin and duplex states as a function of complementary strand concentration. With the finding that these sequences preferentially form hairpins, potential roles these structures could play in regulatory activities are considered.

Self-complementary or palindromic sequences are ubiquitous in DNA and RNA sequences of both prokaryotic and eukaryotic organisms (cf. Gierer, 1966; Rosenberg & Court, 1979; Wells et al., 1980; Gonzalez et al., 1986; Chen et al., 1989). Hairpins or stem−loop structures that form in these sequences have been shown to act as structural vehicles in the regulation of many important genetic processes. For example, hairpin structures are centrally involved in transcription termination where they serve as obvious targets for regulatory proteins (Birchmeier et al., 1983; von Hippel & Yager, 1991, 1992; Jin et al., 1992). In addition, intramolecular structures in both the DNA templates and nascent RNA transcripts act to destabilize transcription complexes and thus are essential for protein independent (intrinsic) termination (Wilson & von Hippel, 1995; Platt, 1986; Yager & von Hippel, 1987). In other studies RNA−RNA interactions between exon regions and ribosomal RNA loops have been used as discriminatory factors for ribosomes to differentiate coding from noncoding regions of mRNA (Mendoza & Lagunez-Otero, 1996). Transcription initiation (Cullen, 1993; Gacy & McMurray, 1994), mRNA stability (Klausner & Harford, 1989; Theil, 1990; Ross & Kobs, 1986; Binder et al., 1989), and several translational processes (Klausner & Harford, 1989; Thiel, 1990; Manzella & Blackshear, 1994) have been shown to involve hairpin structure formation. Recently it has been suggested that intramolecular structures that form in some repeating triplet DNA sequences may be critical structural elements in genetic disease (Mitchell et al., 1995; Mariappan et al., 1996; Sinden & Wells, 1992).

The complete DNA sequence of the chicken glutamine synthetase (GS) reported several years ago (Pu & Young, 1989) revealed, within the 3′ noncoding regions of exon 7,

† This work was supported by NIH Grants GM-39471 (A.S.B.) and EY05063 (A.P.Y.).

* To whom correspondence should be addressed.
‡ University of Illinois.
§ The Ohio State University.

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

S0006-2960(96)01520-6 CCC: $12.00

Table 1: DNA and RNA Sequences Analyzed

DNA	
d(I)	5′-dT T T T T T A A T A A T T A A A A A A-3′
d(II)	5′-dT T T T T T A A T T A T T A A A A A A-3′
d(III)	5′-dA A T A T A T G A A T T C T A A T T A A-3′
d(IV)	5′-dT T A A T T A G A A T T C A T A T A T T-3′
RNA	
r(I)	5′-rU U U U U U A A U A A U U A A A A A A-3′
r(II)	5′-rU U U U U U A A U U A U U A A A A A A-3′

an A·T-rich stretch of approximately 1500 base pairs. From positions 2605 through 2623, a 19-base entirely self-complementary sequence occurs. In this paper, results of gel electrophoretic and optical and calorimetric melting analysis performed on 19-base synthetic DNA and RNA oligomers with sequences of this perfect inverted repeat are reported. The study shows that both the DNA and RNA oligomer strands preferentially form intramolecular hairpins.

MATERIALS AND METHODS

DNA Samples

DNA strands were synthesized on the 1 μmol scale on an automated DNA synthesizer (Applied Biosystems 380B) using the standard β-cyanoethyl phosphoramidite chemistry (Caruthers, 1982). All DNAs terminate in 5′-OH residues. DNA sequences that were prepared are shown in Table 1. The 19-base DNA strand denoted d(I) is referred to as the "primary" DNA strand and corresponds to the coding strand sequence of the glutamine synthetase gene from nucleotides 2605−2623. DNA strand d(II) is the 19-base oligomer complementary to strand d(I). The 20-base DNA strands d(III) and d(IV) have complementary sequences and were used as gel mobility markers. Protocols employed to purify the synthetic DNAs have been published (Paner et al., 1990; Doktycz et al., 1990). Following purification, DNA samples were exhaustively dialyzed vs double-distilled water, vacuum dried, and stored at −20 °C.

RNA Samples

Sequences of the two RNA strands that were prepared are also shown in Table 1. The RNA strand denoted r(I) is the primary RNA strand and corresponds to the sequence of the transcribed mRNA from nucleotide positions 2605−2623. RNA strand r(II) is the exact complement of r(I). Since these strands are the RNA analogs of DNA strands d(I) and d(II), the sequences of strands d(I) and r(II) and d(II) and r(I) are also complementary to one another.

Three independently prepared synthetic batches of RNA sequences r(I) and r(II) were purchased from commercial sources. Two separate RNA syntheses were purchased from The Midland Certified Reagent Co. (Midland, TX). Received as the lyophilized sodium salt, these synthetic RNAs were reported by the supplier to be approximately 95% pure after anion exchange HPLC. The third batch of synthetic RNA strands was purchased from Biosynthesis Incorporated (Lewisville, TX) and received still attached to solid support. RNA strands received from this supplier were cleaved from solid support by treatment with a mixture of 75% ammonia: 25% ethanol and incubated at 60 °C for 4 h. The solution was dried and then deprotected by addition of TBAF (tetrabutylammonium fluoride) and incubation at 20 °C for 16 h. A 1:1 mixture of TEAA (triethylammonium acetate): H_2O was then added, and the solution was dried under vacuum and reconstituted in 100 mM TEAA. This material was then desalted by chromatography on a SepPAK (Waters) C_{18} reverse phase column pre-equilibrated in 2 M TEAA. After RNA solutions were loaded on the column, the column was washed with 100 mM TEAA and H_2O and then eluted with a 1:1 mixture of TEAA:acetonitrile. Recovered material was dried under vacuum and rehydrated in melting buffer (100 mM NaCl, 10 mM sodium phosphate, 1 mM EDTA, pH = 7.0). The total sodium ion concentration of the buffer was 115 mM Na^+.

RNA molecules were purified electrophoretically and recovered by the crush and soak method (Maniatis et al., 1982). Samples were run on 20% polyacrylamide gels containing 8 M urea at 65 °C. Gel bands were back-shadowed with a hand held UV lamp (254 nm), excised, and eluted from gel slices overnight in 1X TBE (50 mM Tris, 50 mM boric acid, 1 mM EDTA, pH = 8.2) containing 300 mM sodium acetate (NAOAc). Purified RNAs were then passed through a Biospin (Biorad) buffer exchange column containing melting buffer, separated into 1 mL aliquots, and strored at −20 °C. Prior to their use in experiments, appropriate amounts of RNA samples were removed and thawed. Once RNA samples were thawed, they were not refrozen.

Optical Melting Curves

To prepare them for optical melting experiments, dried samples of DNA strand d(I) were dissolved and resuspended in 1 mL of melting buffer and diluted with filtered melting buffer to the desired final concentration. For RNA strand r(I) a concentrated aliquot stored in melting buffer was thawed and then diluted in filtered melting buffer to the desired final concentration. Sample and reference solutions were filtered by syringe pressure through 0.45 μm filters. DNA and RNA concentrations of strands d(I) and r(I) were determined using an extinction coefficient of 10 179 M^{-1} cm^{-1} per nucleotide determined from the nearest-neighbor method (Cantor et al., 1970). Melting experiments were performed at DNA strand concentrations over the 53-fold range 0.47−24.8 μM. For the RNA oligomer, strand concentration varied over the 188-fold range 0.3−56.4 μM. Depending on the concentration, optically matched quartz cuvettes with path lengths of 1 or 0.1 cm were employed. The longer path length cuvettes were employed for melting experiments conducted at DNA and RNA strand concentrations $\leq$ 6 μM. For experiments in 1.0 cm path length cuvettes, sample and reference solutions were first filtered, then added to cuvettes, and extensively bubbled with helium. For experiments in 0.1 cm path length cuvettes, sample and reference solutions were filtered several times and then degassed by spinning under vacuum for 1 min in a Speedvac concentrator. These filtered and degassed samples were then carefully added to the cuvettes. For both types of experiments, several drops of mineral oil were lightly layered over the solutions and cuvettes were tightly sealed with Teflon stoppers.

Melting curves of absorbance at 268 nm vs temperature (A_{268} vs T) were collected on a Hewlett-Packard 8450A diode array double-beam spectrophotometer. For experiments using 1 cm path length cuvettes, sample temperature was determined by a temperature probe immersed directly in the sample cuvette. For experiments in the 0.1 cm path length cells, the sample temperature was determined from the cell-

holder temperature. Independent calibration experiments indicated that the solution temperatures of the sample and reference in the 0.1 cm path length cuvettes lag the cell-holder temperature by 3.2 °C. Forward (heating) and reverse (cooling) transitions were collected at a heating−cooling rate of 60 °C/h. Heating and cooling curves were virtually identical at all concentrations, indicating that the melting transitions were entirely reversible. At least four optical melting curves were collected for the DNA and RNA oligomers at each concentration. Data shown are averages obtained from multiple experiments.

Transition curves, A_{268} vs T, were converted to relative absorbance vs T plots by dividing all absorbance readings by the value of the first absorbance reading. Then the transition curves were normalized to upper and lower base lines and converted to θ_B (fraction of broken base pairs) vs T curves. From these curves the differential melting curves, $d\theta_B/dT$ vs T, were constructed. The transition temperature, t_m was determined from the θ_B vs T curves as the temperature where $\theta_B = 0.5$ and from the temperature of the maximum peak height, $(d\theta_B/dT)_{max}$, on differential curves. For unimolecular transitions the t_m values determined from either $\theta_B = 0.5$ or $(d\theta_B/dT)_{max}$ are the same. In general, this is not the case for bimolecular melting transitions. Thermodynamic transition parameters for the hairpin to single-strand transition were obtained from optical melting curves, assuming the transition occurs in a two-state manner. In this graphical analysis of the melting curve (Marky & Breslauer, 1987), the van't Hoff (vH) transition enthalpy is determined from $(d\theta_B/dT)_{max}$,and T_m, viz.,

$$\Delta H_{vH} = 4RT_m^2/(d\theta_B/dT)_{max} \quad (1)$$

The vH entropy is determined directly from the ratio $\Delta H_{vH}/T_m = \Delta S_{vH}$.

Calorimetric Melting Curves

Calorimetric measurements were made on a MC-2 ultrasensitive differential scanning calorimeter (Microcal, Northampton, MA). Samples of the DNA oligomer d(I) for calorimetric measurements were at a concentration (146 μM) 6−310 times higher than for the optical melting curves. Samples of the RNA oligomer r(I) were prepared at 46.5 μM, within the range of concentrations used for the optical melting curves. Calorimetric melting curves were collected as the change in excess heat capacity, ΔC_p, vs temperature, T. After these curves were normalized to total strand concentration, the average buffer base line determined from at least three independent scans of the buffer alone was aligned at the starting temperature and subtracted from these curves. The calorimetric transition enthalpy, ΔH_{cal}, was determined from the integrated area under the base line corrected ΔC_p vs T curve, viz.,

$$\Delta H_{cal} = \int \Delta C_p \, dT \quad (2)$$

The calorimetric transition entropy, ΔS_{cal}, was also determined from the base line corrected ΔC_p as

$$\Delta S_{cal} = \int \Delta C_p/T \, dT \quad (3)$$

At least three forward and reverse ΔC_p vs T scans were made. Reported values of ΔH_{cal} and ΔS_{cal} are the averages from multiple measurements. For all calorimetric melting experiments the net change in heat capacity was nil, i.e., the difference between ΔC_p measured at the initial and final temperatures was zero.

Gel Electrophoresis

Radio Labeling. The primary 19-mer DNA and RNA strands d(I) and r(I) and the 20-mer DNA strands d(III) and d(IV) were labeled on their 5′ ends with ^{32}P and T4 polynucleotide kinase using [γ-^{32}P]ATP. All gel solutions and buffers were prepared with nanopure water that was treated with diethyl pyrocarbonate (DEPC) and autoclaved. The suggested standard procedures were employed (Maniatis et al., 1982). Polynucleotide kinase was purchased from New England Biolabs. Radiolabeled ATP (6000 Ci/mmol, 10 mCi/mL) was purchased from New England Nuclear. Labeled fragments were purified on a 20% denaturing polyacrylamide gel (in the case of the RNA, native gels containing 115 mM NaCl were employed). Corresponding bands were excised, and the oligonucleotides were recovered by the crush and soak method (Maniatis et al., 1982), concentrated in a speed vac concentrator, desalted, and buffer exchanged into melting buffer using a BioSpin-6 (Biorad) spin column. Approximately 70% of the total material contained a radioactive label. Purified labeled samples were stored at 4 °C for no longer than 7 days before use in experiments.

Mixing Experiments. For mixing experiments, 0.1 pmol of labeled primary 19-mer self-complementary strands d(I) and r(I) were mixed with increasing stoichiometric amounts of stock cold (nonlabeled) complementary 19-mer DNA and RNA strands d(II) and r(II) in melting buffer. Mixtures containing ratios of d(I):d(II) and r(I):r(II) at 1:1, 1:5, 1:10, 1:50, 1:100, and 1:1000 were prepared. Chimeric hybrid mixtures were also prepared by mixing d(I) with r(II) or r(I) with d(II) at ratios of 1:10 and 1:1000. Mixtures were made in 15 μL volumes, incubated for 3 min at 90 °C, and slowly cooled overnight at room temperature (20 °C). Three microliters of stock TBE buffer containing 30% glycerol and 0.1% bromphenol blue was added. Mixtures of the 20-mer DNA strands d(III) and d(IV) in nearly equal molar proportions (both 20-mer strands were radio labeled) were also prepared, were treated identically, and served as gel markers. Samples were loaded on 20% polyacrylamide gels prepared from a 30% stock solution (29% acrylamide, 1% bisacrylamide) and electrophoresed at 5 °C, 10 V/cm for 10 h in 1X TBE buffer containing 115 mM NaCl. Gels were autoradiographed using Fuji RX film and intensifying screens at −70 °C.

Autoradiographs were scanned using a Microtek II-XE scanner, and band intensities were quantified by analysis with NIH image I-52 software. Scanned autoradiographs were converted to pixel density (PD) maps using resident software. Bands containing the label were identified as regions with PD higher than that of the background. The relative PD of a particular band provided an estimate of the relative amount of that species present. All PDs were determined from band intensities in the linear response region of the X-ray film. In this procedure relative amounts of material in individual bands were normalized to total material present in all bands in that lane and converted to fraction of total intensity, f_I. For any given experiment, f_I values were reliable to within ±5%. Variation from gel experiment to experiment was slightly greater at ±9%.

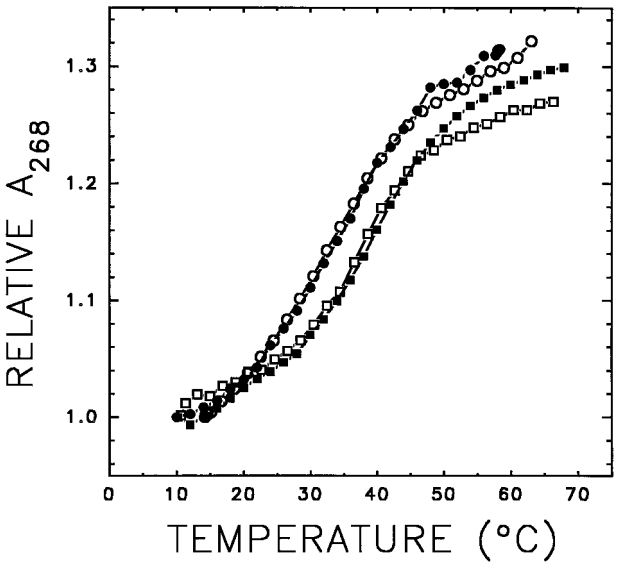

FIGURE 1: Absorbance vs temperature melting curves for the DNA and RNA strands. The relative absorbance changes at 268 nm as a function of temperature for DNA strand d(I) (squares) at concentrations of 0.47 (open symbols) and 24.8 μM (filled symbols) and RNA strand r(I) (circles) at strand concentrations of 0.3 (open symbols) and 56.4 μM (filled symbols). The solvent contained 115 mM Na^+.

RESULTS

Melting Curves of the DNA and RNA Strands

DNA Strand. Results of optical melting experiments for the primary DNA strand d(I) are shown in Figure 1 and Figure 2, panels a and b. In Figure 1 the relative absorbance changes as a function of temperature are shown (squares) for strand d(I) at strand concentrations of 0.47 (open symbols) and 24.8 μM (filled symbols). The average relative hyperchromicity change observed for these transitions was 28.6%. The fraction of broken base pairs, θ_B, vs T collected at the same strand concentrations of 0.47 (solid line) and 24.8 μM (dashed line) of the 19-mer DNA strand are shown in Figure 2, panel a. The differential melting curves, $d\theta_B/dT$ vs T, at these same concentrations are shown in Figure 2, panel b. Spanning the 53-fold concentration range over which melting curves were collected, the observed curves are virtually identical. The t_m values obtained from θ_B vs T curves (at $\theta_B = 0.5$) or the position of the peak height maximum, $(d\theta_B/dT)_{max}$, on differential curves were identical at all DNA concentrations. The average t_m over all concentrations was 37.2 ± 0.4 °C.

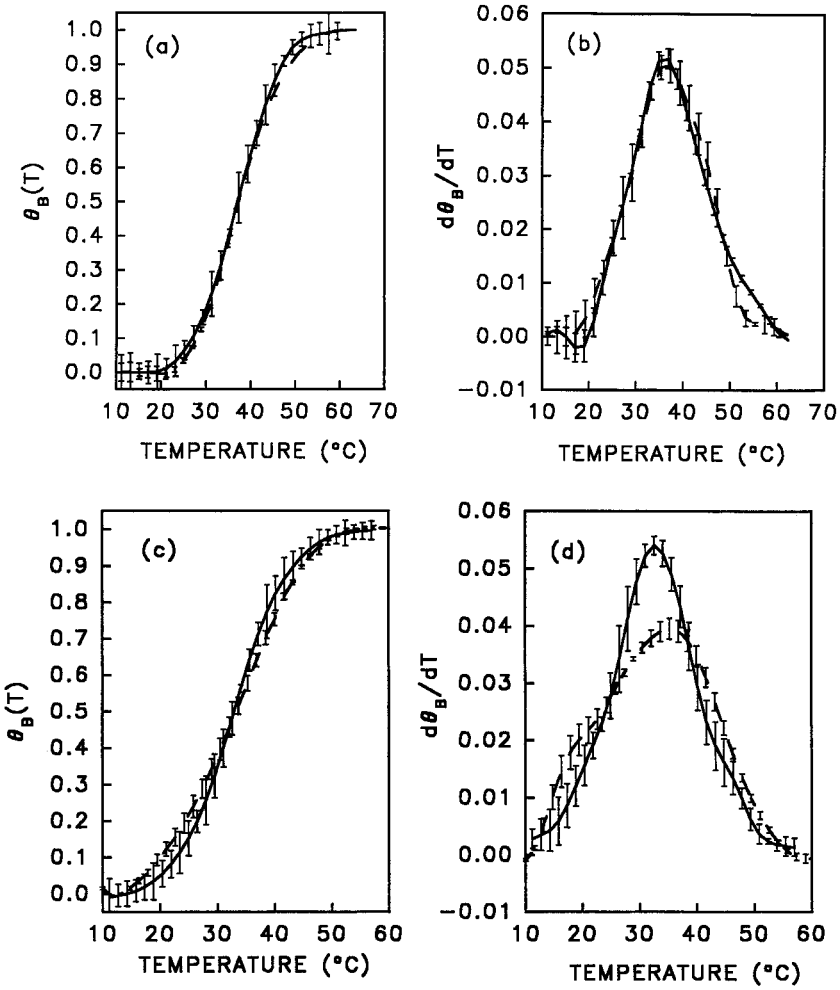

FIGURE 2: Optical melting curves of the 19-base self-complementary DNA and RNA oligomers. (a) The fraction of broken base pairs, $\theta_B(T)$, vs temperature obtained for DNA strand d(I) at DNA concentrations of 0.47 (solid line) and 24.8 μM (broken line) strands in 115 mM Na^+. (b) The differential melting curves, $d\theta_B/dT$ vs temperature obtained from the curves in panel a. Note, the peak positions and curve shapes are essentially identical within the experimental error (indicated by the error bars). The average transition temperature obtained from melting experiments conducted over all strand concentrations was 37.2 ± 0.4 °C. (c) Analogous to panel a except for the RNA strand r(I) at strand concentrations of 0.3 (solid line) and 56.4 μM (broken line). (d) Differential melting curves, $d\theta_B/dT$, vs temperature for r(I) obtained from the curves in panel c. The average transition temperature over all strand concentrations was 32.4 ± 1.1 °C. Note, some differences in the curves shapes are seen at the higher strand concentration, possibly indicative of the minor presence of more than one type of structure that can form in the RNA strand.

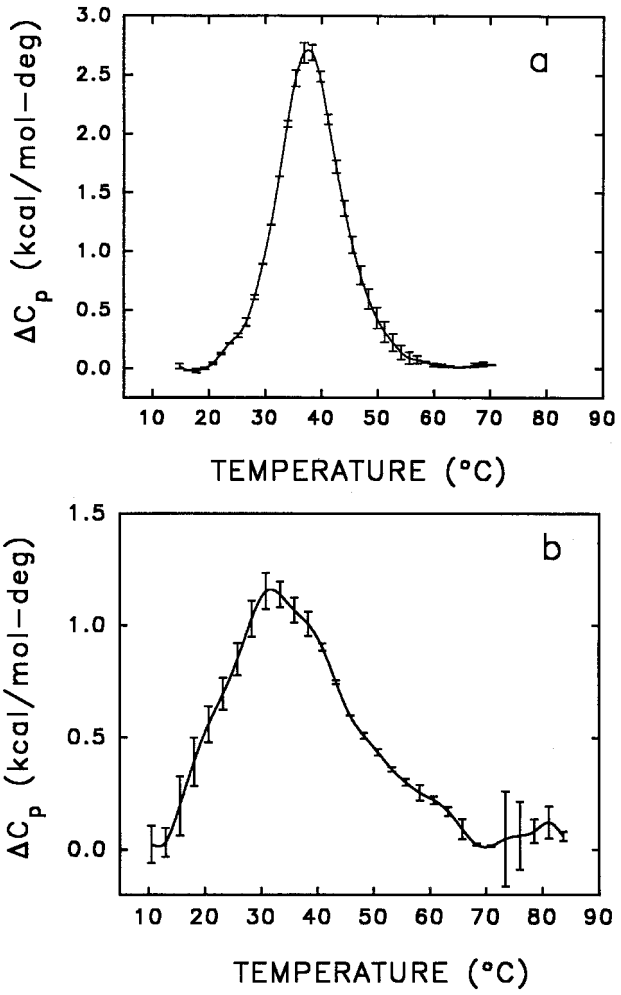

FIGURE 3: Calorimetric melting curves of the 19-base self-complementary DNA and RNA fragments. (a) The average of three independent scans of the excess heat capacity, ΔC_p, vs temperature, for the DNA strand d(I) are shown. These experiments were collected at a strand concentration of 146 μM. The transition temperature (37.6 ± 0.5 °C) is identical to that obtained from the optical experiments shown in Figure 2, panels a and b. (b) The average of three independent scans of ΔC_p vs temperature for the RNA strand r(I). Experiments were conducted at a strand concentration of 46.5 μM. The transition temperature, 32.2 ± 0.5 °C, agrees with optical experiments shown in Figure 2, panels c and d. Error bars denote the experimental reproducibility.

The average calorimetric melting curve, ΔC_p vs T, for d(I) from two independent experiments at a strand concentration of 146 μM is shown in Figure 3, panel a. The transition temperature of the calorimetric curve, $t_m = 37.6 \pm 0.5$ °C, is equivalent to that obtained from optical melting curves even though the calorimetric experiments were conducted at 6−310-fold higher DNA concentrations. The t_m values are summarized in Table 2.

RNA Strand. Results of optical melting experiments for the primary RNA strand r(I) are shown in Figures 1 and 2, panels c and d. The relative absorbance changes as a function of temperature are shown (circles) for strand r(I) at strand concentrations of 0.30 (open symbols) and 56.4 μM (filled symbols). The average relative hyperchromicity change in the r(I) melting transitions was 31.7%, slightly higher than found for d(I) (28.6%). Normalized optical melting curves of r(I) at concentrations of 0.3 (solid line) and 56.4 μM (dashed line) strand r(I) are shown in Figure 2, panel c. The corresponding differential melting curves are shown in Figure 1, panel d. As observed for strand d(I),

Table 2: Thermodynamic Transition Parameters

	strand	
parameter	DNA	RNA
t_m (°C)	37.2 ± 0.4	32.0 ± 0.7
ΔH_{vH} (kcal/mol)	41.6 ± 3.0	33.3 ± 4.2
ΔS_{vH} (cal/K·mol)	134 ± 10	109 ± 14
ΔG_{vH}[a] (kcal/mol)	4.3 ± 0.2	3.0 ± 0.3
$t_{m,cal}$ (°C)	37.6 ± 0.5	32.2 ± 0.5
ΔH_{cal} (kcal/mol)	40.8 ± 0.8	30.7 ± 1.8
ΔS_{cal} (cal/K·mol)	131 ± 1	100 ± 6
ΔG_{cal}[a] (kcal/mol)	4.4 ± 0.5	2.9 ± 0.2

[a] 5 °C.

over the 188-fold range in strand concentration examined, the t_m values were virtually identical at all concentrations of r(I). The average t_m of the optical melting curves was 32.0 ± 0.7 °C.

The average calorimetric melting curve of r(I) obtained from four experiments performed at 46.5 μM is shown in Figure 3, panel b. The average t_m obtained from these experiments is 32.2 ± 0.5 °C, in precise agreement with results from optical melting. Compared to the average calorimetric melting curve of d(I) in Figure 3, panel a, that of r(I) has larger experimental errors and is broader and less symmetric.

As summarized in Table 2, melting temperatures of the primary DNA and RNA strands, d(I) and r(I), are concentration independent over the ranges examined. The transition for strand d(I) is more stable than that for r(I) by 5 °C. Concentration independent melting behavior is a well-known signature of first-order unimolecular melting transitions, as is the case for melting an intramolecular hairpin structure to a nonbonded single strand. While the t_m values of r(I) are concentration independent, the shapes of the differential melting curves shown in Figure 2, panel d, collected at 0.3 and 56.4 μM are clearly different. At the lower concentration (solid line) the differential melting curve is symmetric with a shape similar to that observed for d(I) (Figure 2, panel b). At the higher strand concentration the curve displays a slight shoulder on the low-temperature side of the main transition and is slightly broader with a lower peak height. The shoulder on the optical melting curve of r(I) and asymmetry of its calorimetric melting transition at high strand concentrations probably reveal the presence of a statistically minor population of the bimolecular bulged duplex or hairpin isomers that form in the sequence that are slightly less stable. In addition some end-fraying may occur that results in a slight destabilization of the molecules.

Thermodynamics of the Melting Transitions

The DNA Strand. Thermodynamic transition parameters determined from optical melting and calorimetric melting experiments of d(I) are summarized in Table 2. The transition enthalpy determined from optical melting experiments is $\Delta H_{vH} = 41.6 \pm 3.0$ kcal/mol. The transition entropy is $\Delta S_{vH} = \Delta H_{vH}/T_m = 134 \pm 10$ cal/K·mol (eu). The free energy at 5 °C is $\Delta G_{vH} = \Delta H_{vH} - (278.15)\Delta S_{vH} = 4.3 \pm 0.2$ kcal/mol.

Within experimental error, identical results were obtained from calorimetric melting experiments of d(I). The calorimetrically determined enthalpy, $\Delta H_{cal} = 40.8 \pm 0.8$ kcal/mol, entropy, $\Delta S_{cal} = 131 \pm 1$ eu, and free energy, ΔG_{cal} (5 °C) = 4.4 ± 0.5 kcal/mol are in precise agreement with the transition parameters determined from optical melting curves.

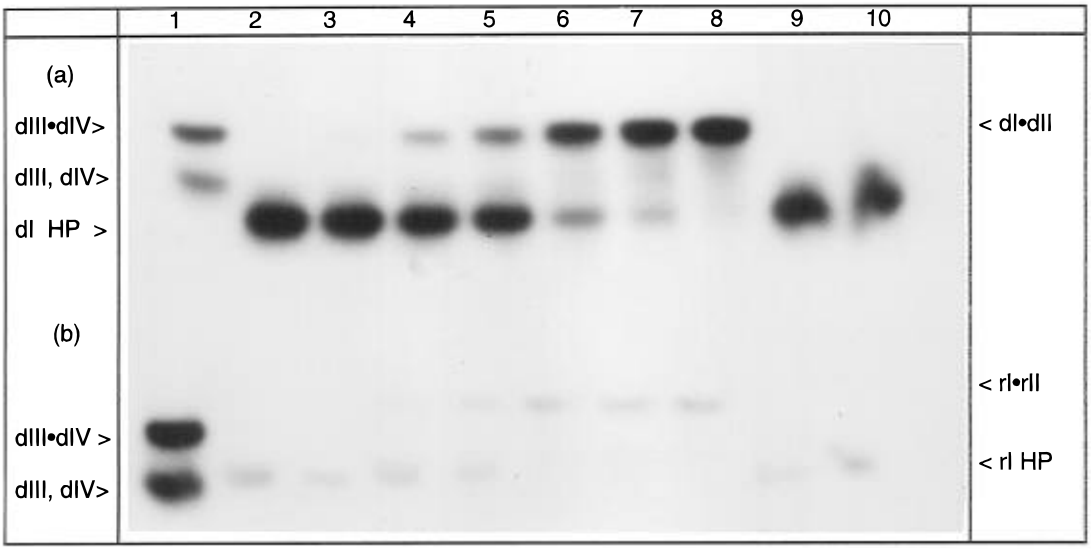

FIGURE 4: Electrophoresis of the DNA and RNA strands and mixtures with their complementary strands on 20%, native polyacrylamide gels. (a) Results for DNA strand d(I) alone and in various mixtures with its complementary DNA strand d(II) or RNA strand r(II). (b) Results for RNA strand r(I) alone and in various mixtures with its complementary RNA strand r(II) or DNA strand d(II). Lanes 1−10 are analogous for both strands and contain the following. Lane 1: DNA markers. As indicated at the left, the upper band corresponds to the duplex formed from DNA strands d(III) and d(IV) (dIII·dIV), the lower band corresponds to a 20-base DNA single strand d(III) or d(IV) (see Figure 1). Lane 2: DNA strand d(I) or r(I) alone. Note, in panel a, lane 2, the DNA strand has greater mobility than the single-strand marker in lane 1, while in panel b the RNA strand has mobility approximately equal to the 20-base single-strand marker. Lane 3: the DNA or RNA strands in a 1:1 stoichiometric mixture with their complementary DNA or RNA strands. Lanes 4−8 contain mixtures as in lane 3, except the stoichiometric ratios were higher at 1:5, 1:10, 1:50, 1:100, and 1:1000, respectively. Lanes 9 and 10 contain the DNA or RNA strands in stoichiometric mixtures at ratios of 1:10 and 1:1000 with their complementary RNA or DNA strands. The slower migrating species seen at excess ratios of the complementary strands correspond to the intermolecular duplex structures.

RNA Strand. The transition thermodynamics obtained from optical melting (the average of 17 experiments on samples obtained from two suppliers) and calorimetric experiments of r(I) are also summarized in Table 2. The van't Hoff transition enthalpy determined from the optical melting curves was $\Delta H_{vH} = 33.3 \pm 4.2$ kcal/mol, the entropy was $\Delta S_{vH} = 109 \pm 14$ eu, and the free energy was $\Delta G_{vH} = 3.0 \pm 0.3$ kcal/mol at 5 °C. Just as found for d(I), the transition thermodynamics of r(I) determined from calorimetry are in close agreement with values determined from optical melting experiments. The calorimetrically determined enthalpy was $\Delta H_{cal} = 30.7 \pm 1.8$ kcal/mol, the entropy was $\Delta S_{cal} = 100 \pm 6$ eu, and the free energy was ΔG_{cal} (5 °C) = 2.9 ± 0.2 kcal/mol.

Such close agreement between the thermodynamic transition parameters derived from the two-state model dependent optical melting analysis and the direct model independent calorimetric measurements reveals two fundamental characteristics of the melting transitions of the secondary structures formed by the 19-base DNA and RNA oligomers. The thermodynamics determined at t_m for the melting transitions for both molecules are concentration independent and thus are consistent with transitions of unimolecular hairpins to nonbonded single strands. In addition, the precise agreement between the van't Hoff and calorimetric thermodynamics concur that, within the level of experimental error, the melting transitions of these structures cannot be distinguished from the two-state model prediction. For the reasons mentioned earlier, the slight shoulder on the differential melting curve of r(I), seen at high strand concentrations, may represent an exception to this. Even so, deviations from two-state behavior, as manifested in the melting curve shapes, must be relatively small because the analysis procedure is insensitive to their presence.

Gel Electrophoresis Reveals Unimolecular Hairpins Form from the DNA and RNA Strands

Results of gel electrophoresis of strands d(I) and r(I) alone and in mixtures with their DNA or RNA complementary strands are shown in Figure 4, panels a and b. Figure 4, panel a, displays results of mixing experiments for d(I). Results for r(I) are shown in Figure 4, panel b. Although the band intensities on the RNA gel panel b are less intense than those of the DNA experiments in panel a (due to lower activity of the radioactive label), the major trends are quite evident. In both gels, the first lane contains the markers and shows two bands, one corresponding to a 20-base single strand (d(III) or d(IV)) and the duplex having one each of strands d(III) and d(IV). Two bands result because d(III) and d(IV) were not mixed in exactly equal stoichiometric amounts, leaving one strand in excess of the other. Lane 2 shows results for strand d(I) or r(I) alone. As seen in panel a, d(I) has mobility greater than the 20-base pair duplex or 20-base single strands in lane 1. In contrast, in panel b under the same conditions, r(I) has mobility equivalent to the 20-base single-strand d(III) or d(IV). Results of incubating mixtures of the labeled d(I) and r(I) strands with increasing amounts of their complementary strands d(II) and r(II) are shown in lanes 3−8 of Figure 4, panels a and b. Results of mixtures of d(I) with r(II) and r(I) with d(II) are shown in lanes 9 and 10.

At a ratio of 1:1 of labeled d(I) or r(I) with their complements, (lane 3) 100% of the material migrates to the same position as the labeled strand alone (lane 2). As clearly shown in lanes 4−8, with increasing amounts of the complementary strand, increasing amounts of a lower mobility species with migration distance of the 20-mer duplex (presumed to be the 19-base pair bimolecular duplex) are seen in addition to the band observed in lane 2 for d(I) or

r(I) alone. For d(I) and r(I), conversion of the faster mobility to the slower mobility species does not occur until a stoichiometric ratio of d(I):d(II) or r(I):r(II) = 1:1000 is reached (lane 8). Although this is not clearly evident on the photograph in Figure 4, panel b, the digital scan of the gel (not shown) revealed that conversion of the faster mobility species for r(I) was not complete until a ratio of r(I):r(II) = 1:1000.

These results indicate that the DNA and RNA strands, d(I) and r(I), alone at equilibrium preferentially form intramolecular structures. In excess stoichiometric mixtures with their complement strands, the equilibrium is shifted to the bimolecular DNA/DNA and RNA/RNA duplex forms. In contrast, for the hybrid mixtures d(I):r(II) or r(I):d(II) at comparable strand ratios (1:10 and 1:1000, lanes 9−10 in Figure 4), no evidence for formation of the DNA/RNA hybrid duplexes was seen. Quickly cooling the DNA or RNA samples did not effect the results (data not shown), indicating we have examined the equilibrium (not kinetic intermediate) structures formed by the strands.

DNA and RNA strands d(I), d(II), r(I), and r(II) were each electrophoresed individually under native conditions (see Materials and Methods) at strand concentrations from 30 to 150 μM. These concentrations span the range of those used in the optical and calorimetric melting studies. Results of this analysis indicated that the DNA and RNA strands have the same gel mobility at all strand concentrations (not shown). This is consistent with the agreement in the thermodynamic parameters evaluated at widely different concentrations by optical and calorimetric methods, i.e., these strands form unimolecular hairpins at all concentrations examined.

Multiple preparations of labeled RNA stocks were often needed to ensure highly pure RNA absent of degradation products induced by contaminants or caused by the radioactive label itself. We observed, over time, differences in band intensities of labeled DNA and RNA prepared from the same stock of [γ-^{32}P]ATP. Comparison of the mobilities of the DNA and RNA strands alone (lanes 2 in Figure 4, panels a and b, respectively) reveals that d(I) has a greater mobility than the 20-base pair duplex and 20 base single strands while the analogous RNA single-strand r(I) runs with mobility between those of the 20-base single strand and 20 base pair duplex. Clearly, under nondenaturing conditions the RNA and DNA strands (r(I) and d(I)) display very different gel mobilities. Under denaturing (8 M urea, 60 °C), but otherwise identical, electrophoretic conditions, mobilities of d(I) and r(I) were indistinguishable (data not shown). The mobility differences for d(I) and r(I) observed under nondenaturing conditions are likely to be due to different base contents (uridine vs thymine) and/or differences in the sizes of the stem and loops in the hairpins that form from d(I) and r(I) that effect their gel electrophoretic mobilities differently.

Digital scans and analysis of the electrophoretograms enabled semiquantitative thermodynamic analysis of the mixing equilibria. Following our interpretations of the mobility positions of the DNA and RNA species on the electrophoretograms, fractional intensities of the faster mobility and slower mobility bands presumably correspond to the fraction of labeled strands in the hairpin, F_{HP*}, and the fraction of labeled strands in the duplex, F_{D*} ($F_{HP*} + F_{D*} = 1$). These fractions are plotted as a function of complement strand stoichiometry (Sc:S*) in Figure 5.

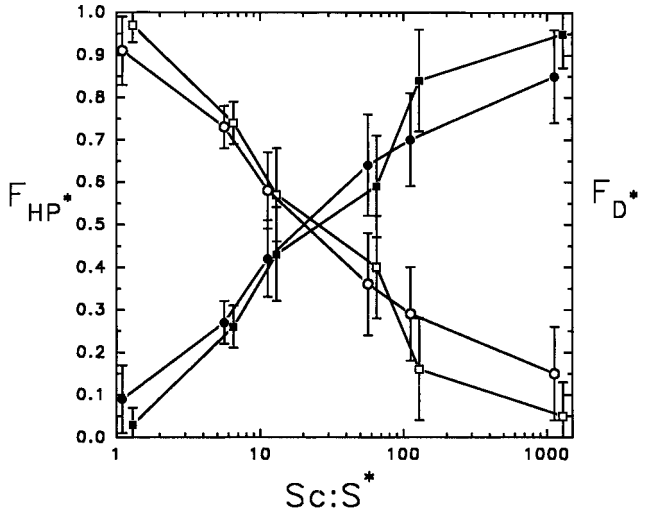

FIGURE 5: Relative fractions of hairpin and duplex as a function of complementary strand concentration. The fraction of hairpin, F_{HP*}, and duplex, F_{D*}, determined from the digitally scanned fractional intensities of the respective bands on the autoradiograms shown in Figure 3 as a function of stoichiometric ratio of the primary DNA (d(I)) or RNA strand (r(I)) and its complementary DNA (d(II)) or RNA strand (r(II)), Sc:S* (see text). Results for the DNA are squares, those for the RNA are circles. Open symbols depict F_H* and filled symbols F_D*. The error bars denote the deviations from the average for three independent mixing experiments like those depicted in Figure 3.

Results depicted are the averages of at least three separate independent experiments. Clearly, at this level of resolution the DNA and RNA strands respond similarly in mixtures with their complimentary DNA and RNA strands. Quantitative insight into this behavior can be gained by considering the coupled equilibrium reaction in which self-complementary DNA or RNA strands can self-fold to form intramolecular hairpins or associate with each other to form a duplex. The representative reactions are

$$\mathrm{HP^* + HPc} \overset{K_1}{\rightleftharpoons} \mathrm{S^* + Sc} \overset{K_2}{\rightleftharpoons} \mathrm{D^*}$$

where $K_1 = K_{HP*}K_{HPc}$, with K_{HP*} and K_{HPc} the equilibrium constants for hairpin formation from each of the individual single strands S* and Sc, and K_2 is the equilibrium constant for duplex formation containing one each of the individual strands. The asterisks denote the labeled strand or duplex complex containing one labeled strand (S* is either d(I) or r(I)). In mixing experiments, S* (d(I) or r(I)) was fixed and its complementary strand Sc(d(II) or r(II)) was varied. The quantity of interest is the ratio of the equilibrium constants for hairpin and duplex formation as a function of added complementary strand concentration $[Sc]_T$,

Experimentally, from gel analysis we measured F_{HP*} and F_{D*}, the fractions of labeled material in the hairpin and duplex bands, respectively. In terms of these quantities and the total concentrations, $[S^*]_T$ and $[Sc]_T$,

$$K_1/K_2 = (F_{HP*}/F_{D*})([Sc]_T - F_{D*}[S^*]_T)$$

The free energy difference between the hairpin and duplex states for the labeled single strand is

$$\Delta\Delta G = -RT \ln(K_1/K_2) = \Delta G_{HP*} - \Delta G_D \qquad (5)$$

where ΔG_{HP*} and ΔG_D are the free energies of formation for the hairpin and duplex states. Plots of $\Delta\Delta G$ (5 °C) vs stoichiometric strand ratios Sc:S* for S* = d(I), Sc = d(II)

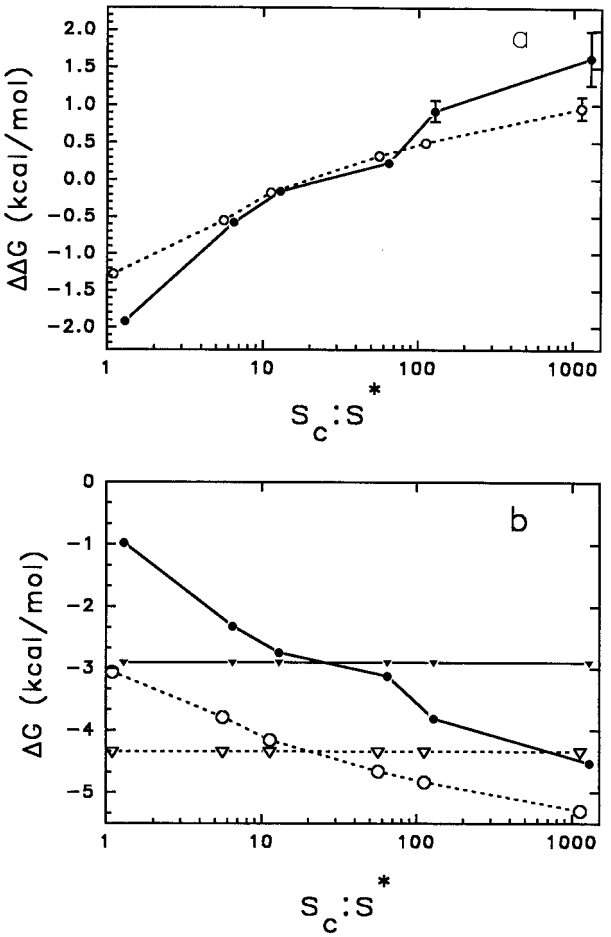

FIGURE 6: Relative free energies of the duplex and hairpin states as a function of strand stoichiometric ratio. (a) Plots of the free energy difference, $\Delta\Delta G = \Delta G_{HP^*} - \Delta G_D$, for the primary DNA strand (solid line) and the primary RNA strand (broken line) vs complementary strand stoichiometry (Sc:S*). Clearly, at lower Sc:S* ratios the hairpin is more favored than the duplex ($\Delta\Delta G \leq 0$), and this remains the case until $\Delta\Delta G \approx 0$ when Sc:S* $\geq$ 30:1. (b) Free energy of the duplex formed with the primary DNA strand d(I) and its complement (solid line) and the primary strand r(I) with its complement (broken line) vs Sc:S*. The free energy of hairpin formation for these sequences, which are independent of the Sc:S* are shown as the inverted triangles.

(solid line) or S* = r(I), Sc = r(II) (broken line) are shown in Figure 6, panel a. From an independent melting experiment of strand S* alone (d(I) or r(I); Figure 2 and 3 and Table 2), the free energies of hairpin formation, ΔG_{HP^*}, were evaluated. With this complementary data and employing eq 5 with the results presented in Figure 6, panel a, the free energy of duplex formation as a function of strand stoichiometric ratio Sc:S* can be determined. As shown in Figure 6, panel b, plots of ΔG_D vs Sc:S* for d(I) (solid line) and r(I) (broken line) decrease with increasing values of Sc:S*. The free energy of the hairpins, ΔG_{HP^*}, which are independent of Sc:S*, are also shown as horizontal lines on the plot in Figure 6, panel b. Clearly, for both d(I) and r(I) the hairpin is much more favorable than the duplex until Sc:S* $\geq$ 30:1.

DISCUSSION

Results of optical and calorimetric melting and gel electrophoretic analysis of mixing reactions for the 19-base DNA and RNA inverted repeat sequences corresponding to positions 2605−2623 of the GS gene showed that these naturally occurring sequences preferentially form intramolecular hairpins. The question arises as to the relevance of these findings with respect to the situation in the actual gene environment where short self-complementary regions are flanked on either side by noncomplementary sequences. There is some evidence that even though palindromic sequences are flanked by noncomplementary regions, they may be capable of inducing intramolecular structure formation. Melting studies of DNA fragments 110−205 base pairs in length containing 19−25 base pair palindromic regions imbedded in longer (nonpalindromic) sequences found evidence for intramolecular structures during their melting transitions (McCampbell et al., 1989). Therefore, it may not be unreasonable to expect that the 19-base partially self-complementary strands studied here could similarly form intramolecular structures when imbedded in longer sequences. Further studies will be required to verify that this is actually the case.

Positioning of these hairpin structures within the 3′ noncoding region suggests that they may play a role in mRNA processing or in regulating mRNA stability. The sequence is located approximately 100 base pairs upstream of the GS polyadenylation site (Pu & Young, 1989). Since the GS gene does not contain a perfect version of the canonical endonucleolytic/polyadenylation signal AATAA (Proudfoot & Brownlee, 1976), it is possible that hairpin formation in the DNA may play a role in the formation of a mature 3′ end. The inability of the RNA and DNA strands to form a hybrid duplex may indicate that these hairpins are formed quickly in the transcription process and may help to destabilize the transcription complex, thereby effecting termination.

In the embryonic chicken retina, GS RNA has a relatively short half-life of approximately 1.5 h. Moreover, enhanced expression of GS during retinal development might involve alterations in GS mRNA stability (Patejunas & Young, 1990). It is intriguing to note that 5′-A-A-U-3′ sequences positioned within loops in a secondary structure of apolipoprotein II mRNA are specific target sites for mRNA breakdown (Binder et al., 1989). Coincidently, the stem−loop structure formed from RNA strand r(I) would probably contain the sequence 5′-A-A-U-A-A-U-A-3′ (or portions of it) within the loop. A hairpin structure formed by nucleotides 2605−2623 might help confer instability of GS mRNA, possibly via a developmentally regulated process.

Determining whether the stem−loop structures described above actually participate in regulating GS gene expression will clearly require further direct investigations. Having demonstrated in this initial physical study that DNA and RNA oligomers with sequences of nucleotides 2605−2623 can form hairpin structures, the impetus is provided for functional studies aimed at ascertaining the biological role of these structures.

REFERENCES

Binder, R., Hwang, S.-P. L., Ratnasabapathy, R., & Williams, D. L. (1989) *J. Biol. Chem. 264*, 16910−16918.

Birchmeier, C,, Folk, W., & Birnstile, M. L. (1983) *Cell 35*, 433−440.

Cantor, C. R., Warshaw, M. M., & Shapiro, H. (1970) *Biopolymers 9*, 1059−1077.

Caruthers, M. H. (1982) in *Chemical and Enzymatic Synthesis of Gene Fragments* (Gassen, H. G., & Lang, A., Eds.) pp 71−79, Verlag-Chemie, Weinheim, FRG.

Chen, K. C., Tyson, J. J., Lederman, M., Stoudt, E. R., & Bates, R. C. (1989) *J. Mol. Biol. 208*, 283−296.

Cullen, B. R. (1993) *Cell 73*, 417−420.

Doktycz, M. J., Paner, T. M., Amaratunga, M., & Benight, A. S. (1990) *Biopolymers 30*, 829−845.
Gacy, A. M., & McMurray, C. T. (1994) *Biochemistry 33*, 11951−11959.
Gierer, A. (1966) *Nature (London) 212*, 1480−1481.
Gonzalez, A., Talavera, A., Almendral, J. M., & Vinela, E. (1986) *Nucleic Acids Res. 14*, 6835−6844.
Jin, D., Burgess, R. R., Richardson, J. P., & Gross, C. (1991) *Proc. Natl. Acad. Sci. U.S.A. 89*, 1453−1457.
Klausner, R. D., & Harford, J. B. (1989) *Science 246*, 870−873.
Maniatis, T., Fritsch, E. F., & Sambrook, J. (1982) *Molecular Cloning-A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, pp 122−123.
Manzella, J. M., & Blackshear, P. J. (1990) *J. Biol. Chem. 265*, 11817−11822.
Mariappan, S. V. S., Catasti, P., Chen, X., Ratliff, R., Moyzis, R. K., Bradbury, E. M., & Gupta, G. (1996) *Nucleic Acids Res. 24*, 784−792.
Marky, L. A., & Breslauer, K. J. (1987) *Biopolymers 26*, 1601−1620.
McCampbell, C. R., Wartell, R. M., & Plaskon, R. R. (1989) *Biopolymers 28*, 1745−1758.
Mendoza, L., & Lagunez-Otero, J. (1996) *Biosystems 38*, 45−49.
Mitchell, J. E., Newbury, S. F., & McClellan, J. A. (1995) *Nucleic Acids Res. 23*, 1876−1881.
Paner, T. M., Amaratunga, M., Doktycz, M. J., & Benight, A. S. (1990) *Biopolymers 29*, 1715−1734.
Patejunas, G., & Young, A. P. (1990) *J. Biol. Chem. 265*, 15280−15285.
Platt, T. (1986) *Annu. Rev. Biochem. 55*, 339−372.
Proudfoot, N. J., & Brownlee, G. G. (1976) *Nature 263*, 211−214.
Pu, H., & Young, A. P. (1989) *Gene 81*, 169−175.
Rosenberg, M., & Court, D. (1979) *Annu. Rev. Genet. 13*, 319−351.
Ross, J., & Kobs, G. (1986) *J. Mol. Biol. 188*, 579−593.
Sinden, R. R., & Wells, R. D. (1992) *Curr. Opin. Biotechnol. 3*, 612−619.
Thiel, E. C. (1990) *J. Biol. Chem. 265*, 4771−4774.
von Hippel, P. H., & Yager, T. D. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 2307−2311.
von Hippel, P. H., & Yager, T. D. (1992) *Science 255*, 809−812.
Wells, R. D., Goodman, T. C., Hillen, W., Horn, G. T., Klein, R. D., Larson, J. E., Muller, U. R., Neuendorf, S. K., Panayotatos, N., & Stirdivant, S. M. (1980) *Prog. Nucleic Acid Res. Mol. Biol. 25*, 167−267.
Wilson, K. S., & von Hippel, P. H. (1995) *Proc. Natl. Acad. Sci. U.S.A. 92*, 8793−8797.
Yager, T. D., & von Hippel, P. H. (1987) in *Escherichia coli and Salmonella typhimurium: Cellular and Molecular Biology* ed. (Neidhardt, F. C., Ed.) Am. Soc. Microbiol., Washington, DC, pp 1241−1275.

BI9615203