A Synthetic Model for Triple-Helical Domains in Self-Splicing Group I Introns Studied by Ultraviolet and Circular Dichroism Spectroscopy†

Munna Sarkar,‡ Susannah Sigurdsson,§ Sebastian Tomac,‡ Srikanta Sen,‡ Eriks Rozners,§ Britt-Marie Sjöberg,|| Roger Strömberg,§,⊥ and Astrid Gräslund*,‡

Departments of Biophysics, Organic Chemistry, and Molecular Biology, Arrhenius Laboratories, Stockholm University, S-10691 Stockholm, Sweden

Received October 2, 1995; Revised Manuscript Received January 5, 1996⊗

ABSTRACT: Structural studies were performed on synthetic oligonucleotides with sequences corresponding to the P4/P6 and J3/4, J6/7 regions of the self-splicing group I intron of the bacteriophage T4 *nrdB* pre-mRNA, which correspond to the proposed *triple-helical domain* in the *Tetrahymena thermophila* intron. A 23-mer RNA was synthesized as a mixed ribo−deoxyribo oligonucleotide, modeling an expected base-paired region P4 along with the J3/4 and P6 (5′-end bases of P6) regions. A third strand modeling the 3′-end bases of P6 and J6/7 regions, with which a triple helix may form, was synthesized as a pure oligoribonucleotide (7-mer RNA). The interactions of these oligonucleotides have been characterized by UV and circular dichroism (CD) spectroscopy. The results show that the 23-mer RNA forms a stable hairpin modeling the P4 base-paired region. Triple helix association between the 23-mer RNA hairpin and the 7-mer RNA single strand was detected by CD in the presence of Mg^{2+} (>5 mM) but not in presence of a monovalent cation like Na^{+} (up to 500 mM). Studies on selected variants of both 7-mer and 23-mer RNAs were carried out. The results show that for the association of the two partner strands not only the formation of P6 helix but also triplet interactions between the two strands are required. The association of the two strands in general follow a pattern predicted by comparative sequence analysis. Parallel studies on pure oligodeoxyribonucleotides having base sequences corresponding to those of the oligoribonucleotides showed no evidence of association under similar conditions, which could indicate that the 2′-hydroxyl groups of the riboses might play an important role in hydrogen bonding to form the required nucleoside triples. Molecular modeling studies on the proposed "plaited triple helix" formed by the association of the 23-mer RNA hairpin and 7-mer RNA single strand showed that the structure is sterically and energetically feasible.

The three-dimensional structures of group I intron RNAs determine their function in catalyzing the transesterification reactions of self-splicing. Secondary structure models for this class of catalytic RNA were derived by comparative sequence analysis (Davies et al., 1982; Michel et al., 1982). Many of these proposed structures have been confirmed and studied in detail by site-directed mutagenesis and chemical modification studies (Couture et al., 1990; Inoue & Cech, 1985; Christian & Yarus, 1993; Wang & Cech, 1994). Tertiary interactions between secondary structural elements play a major role in determining the three-dimensional structures in RNA molecules. Systematic comparative sequence analysis has revealed a number of potential long-range interactions (Michel et al., 1989, 1990; Michel & Westhof, 1990). This has led to the proposal of a three-dimensional structure for the catalytic core of the group I intron (Michel & Westhof, 1990). Several of its basic features have been supported by experimental evidence [cf. Cech et al. (1994), Celander and Cech (1991), Pyle et al. (1992), Green and Szostak (1994), Wang and Cech (1992), and Murphy et al. (1994)].

Nucleoside triple interactions involving major and minor grooves have been proposed between paired regions P4 and P6 and single-stranded joining regions J6/7 and J3/4, respectively, by observing covariations in phylogeny (Michel et al. 1989, 1990; Michel & Westhof, 1990). Model building, based on the most common nucleotide combinations found at these positions, led to the proposal that the interactions involved the formation of base triples with the third strand J6/7, interacting in the major groove of the P4 helix (Michel & Westhof, 1990). The existence of these major groove base triples was supported by compensatory mutational analysis in the *sunY* and *Tetrahymena* introns (Michel et al., 1990) and also by comparative sequence analysis of *in vitro* genetically selected variants (Green & Szostak, 1994). The base triples between P6 and J3/4 were accommodated in the model proposed by Michel and Westhof (1990) by invoking interactions on the minor groove side of the P6 base pairs. These P4-J6/7 and P6-J3/4 nucleoside triple interactions lead to the formation of a so-called "*triple-helical domain*" which might be important in guiding the proposed coaxial orientation of the P4 and P6 helices with respect to P7 and may act as a nucleus around which the remaining peripheral structural domain assembles

† This study was supported by grants from the Swedish Natural Science Research Council, the Swedish Institute, Procordia Research Funds, the Computational Fund at the University of Umeå, and the Magnus Bergwall's Foundation.

* To whom correspondence should be addressed.
‡ Department of Biophysics.
§ Department of Organic Chemistry.
|| Department of Molecular Biology.
⊥ Present address: Laboratory of Organic and Bioorganic Chemistry, Department of Medical Biochemistry and Biophysics, Karolinska Institutet, S-17177 Stockholm, Sweden.

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

0006-2960/96/0435-4678$12.00/0

in the group I intron from *Tetrahymena thermophila* (Michel & Westhof, 1990). In a recent study (Doudna & Cech, 1995) it has been specifically shown that these nucleotides, which are proposed to engage in triple strand interactions with stems P4 and P6, are involved in tertiary interactions between the domains of the introns.

A simple schematic model for the formation of nucleoside triples in the P4−P6 junction was put forward by Chastain and Tinoco (1992a, 1993) on the basis of a pseudoknot structure (Pleij et al., 1985). In this model the P4−P6 helices are rotated, so that they can accommodate the J3/4 strand in the minor groove of P6 and the J6/7 strand in the major groove of P4. Their NMR results on a single oligoribonucleotide that models the P6 helix along with a part of the P4 helix and J3/4 single-stranded region show that it is possible to accommodate the J3/4 single strand in the minor groove of the P6 helix, forming nucleoside triples in which interactions through ribose hydroxyls play an important role.

The present study concerns oligonucleotides that could model the tertiary interactions in a triple helical domain at the active site of the self-splicing group I intron from the bacteriophage T4 *nrdB* pre-mRNA, which has a core sequence analogy (Sjöberg et al., 1986) with the *Tetrahymena* system. Our model system consists of a heptaribonucleotide (7-mer RNA) containing the 3′-end bases of P6 and the J6/7 sequence (nucleotides 449−455) and a 23-mer mixed ribo−deoxyribo oligonucleotide (23-mer RNA) containing the J3/4-P4 (5′-end bases of P4) sequence (nucleotides 42−51), followed by three deoxythymidines (3dT) replacing the P5 loop and the P4−P6 (3′-end bases of P4 and 5′-end bases of P6, respectively) sequence (nucleotides 70−79). The P5 loop in the original sequence was substituted by 3dT, because they are expected to form a stable hairpin loop (Baxter et al., 1993; Rentzeperis et al., 1993) that should not be involved in RNA tertiary interactions. The nucleotide numbers used in the text all correspond to those in the original intron sequence. Figure 1a shows the relevant parts of the pre-mRNA based on previously suggested models (Michel & Westhof, 1990; Chastain & Tinoco, 1992a, 1993). Figure 1b shows the synthesized sequences used for this study. A new aspect of the present model system of the triple helical domain, compared to the previous *Tetrahymena* model system (Chastain & Tinoco, 1992a, 1993), is that it is composed of two separate molecules each with consecutive sequences according to the native one that model not only the P4−P6 helices and J3/4 single-stranded region but also the J6/7 region. The assembly process, therefore, in certain aspects might resemble the native folding. If the previously proposed nucleoside triple interactions between P4−P6 helices and J6/7 and J3/4 regions are incorporated in our sequences (Michel et al., 1989, 1990; Michel & Westhof, 1990; Green & Szostak, 1994; Chastain & Tinoco, 1992a, 1993), then the two strands (7-mer and 23-mer RNA) of our model system are expected to associate as shown in Figure 1c.

The principal aim of the present study was to investigate how our model system mimics the formation and structure of the triple helical domain of the intron. We have mainly used UV and circular dichroism spectroscopy to study the synthetic model system and its sequence-dependent structure. First we showed that the 23-mer RNA molecule forms a stable hairpin modeling the P4 base-paired region. Second, we investigated the conditions that lead to the association between the 23-mer RNA hairpin and the 7-mer RNA.

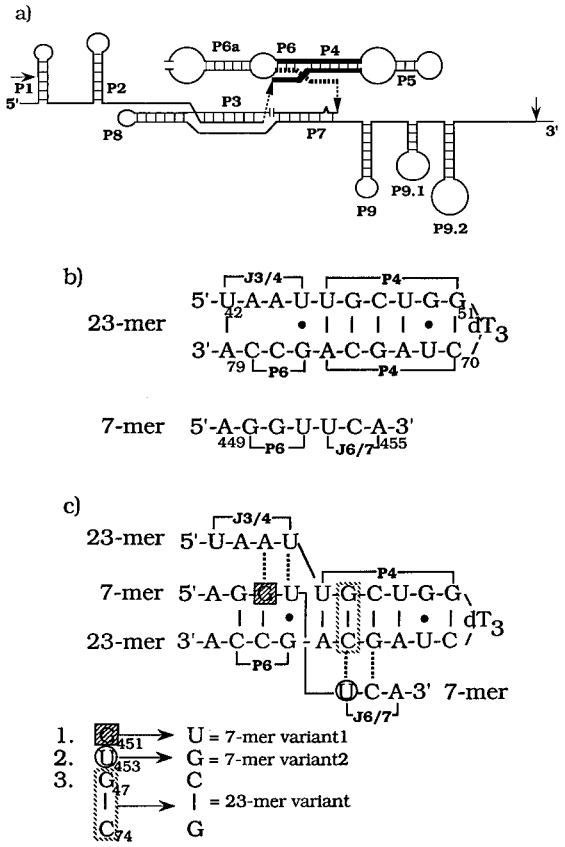

FIGURE 1: (a) Schematic representation of the intron of bacteriophage T4 *nrdB* pre-mRNA. The thick solid and the dotted lines represent the regions modeled by 23-mer RNA and 7-mer RNA, respectively. The dotted lines with arrowheads show the continuation of the base sequence. The arrows represent the 5′ and 3′ splice sites. (b) Model sequences of 23-mer RNA (nucleotides 42−51 and 70−79; P5 loop exchanged for 3dT, i.e., for nucleotides 52−69) and 7-mer RNA (nucleotides 449−455). (c) Model sequences with tertiary structure modeling the coaxial P4 and P6 helices (Murphy et al., 1994) and suggested triplet interactions indicated by dotted bonds [cf. Michel and Westhof (1990), and Michel et al. (1990)]. The bases that have been substituted in the different variants of 7-mer RNA and 23-mer RNA are marked as shown.

Studies on selected variants (Figure 1c) of both 23-mer RNA and 7-mer RNA were carried out to highlight the specific contributions of the position and the nature of the bases in this kind of triple helical association and also to probe if the model may account for certain phylogenetic observations of covariation of nucleotides. To probe the importance of the 2′-hydroxyl groups for the formation of the structures, we have also studied the corresponding deoxyribonucleotide sequences (23-mer DNA and 7-mer DNA).

MATERIALS AND METHODS

Synthesis of Oligonucleotides. The mixed ribo−deoxyribo oligonucleotide (23-mer RNA) and the oligoribonucleotide (7-mer RNA) were synthesized using the H-phosphonate method for RNA synthesis (Garegg et al., 1986b). The procedure used here is an improved version (Westman et al., 1994; E. Rozners and R. Strömberg, unpublished results) where the heterocyclic bases of the building blocks are N-protected with more labile acyl groups (Westman et al., 1993) in order to avoid partial loss of silyl groups and some cleavage of oligomers that can occur if deprotection of the bases requires harsh ammoniolytic conditions (Stawinski et al., 1988).

The solid support used for oligonucleotide synthesis was long-chain alkylamine-controlled pore glass (LCAA-CPG from Pierce) functionalized with 5′-*O*-(4-monomethoxytrityl)-2′- or 3′-*O*-(2-chlorobenzoyl)adenosine 2′- or 3′-*O*-succinates (∼20−28 μmol/g) (Rozners et al., 1995) using standard procedures (Atkinson et al., 1984). The automated synthesizer was a modified Gene Assembler (Pharmacia) and synthetic steps (detritylation with 1% trifluoroacetic acid in dichloroethane for 1 min, condensation during 1.5 min using the 30 mM H-phosphonate building block and 90 mM pivaloyl chloride in pyridine−acetonitrile (1:3 v/v), and oxidation with 2% iodine in pyridine−water (98:2) for 30 min) were as in a recent study of 2′-protection (Rozners et al., 1994).

The support was, after completion of the cycles in the automated synthesizer, transferred to a 2 mL cryovial (with screw cap). For cleavage of the oligomers from the support and deprotection of the bases, 32% concentrated ammonia (aqueous)−ethanol (3:1, 400 μL) was added and the mixture left for 14−16 h at room temperature. The ammonia solutions were filtered from the LCAA-CPG (using a Millex-GV13 filter, 0.22 μm, 13 mm), and the supports were washed with an equal volume of ammonia−ethanol that was pooled together with the first one. The ammonia−ethanol solutions were evaporated under reduced pressure, the residues redissolved in distilled water, and the resulting solutions lyophilized.

Subsequent removal of the 2′-*O*-*tert*-butyldimethylsilyl protecting groups was done with the reagent triethylamine trihydrofluoride (Gasparutto et al., 1992) that has been shown to be more reliable (Westman & Strömberg, 1994) than the commonly used tetrabutylammonium fluoride. The residue obtained when lyophilizing after ammoniolysis was dissolved in triethylamine trihydrofluoride (300 μL) and left in the reagent for 14−16 h at room temperature. Water (300 μL) was subsequently added, and the solution was extracted five times with ethyl acetate (5 × 500 μL) and then lyophilized prior to HPLC purification.

The first purification step for the oligonucleotides (except the 23-mer variant and 7-mer RNA variant 2, see below) was done by anion-exchange HPLC (on a Bakerbond Widepore PEI column, 5 μm, 4.6 × 250 mm) using a linear gradient from 1.0 μM to 0.3 M phosphate buffer (pH = 6.5) in MeCN−water (3:7) in 90 min and a flow rate of 1 mL/min. The HPLC column and gradient used for the 23-mer RNA variant was a Shandon Hypersil WPSAX (5 μm, 7 × 200 mm) and 50−250 mM $LiClO_4$ in 20 mM sodium acetate (pH 6.5, 30% acetonitrile, 2.5 mL/min) for 60 min. A Dionex Nucleo Pac PA-100 column (4 × 250 mm) with 0−60 mM $LiClO_4$ in 20 mM sodium acetate (pH 6.5, 10% acetonitrile, 1 mL/min) for 25 min was used for purification of the 7-mer variant 2. The oligonucleotides were dissolved in 0.5 mL of 1.0 mM buffer solution and filtered through a disposable C-18 cartridge (Waters Sep Pac). The cartridge was washed with another 0.5 mL of buffer, and the combined fractions were filtered through a disposable syringe filter (Millex-GV13 filter, 0.22 μm, 13 mm) before injection. The collected fractions were lyophilized before the next purification step.

Further purification of the oligonucleotides was done by reversed-phase HPLC (on a Supelcosil LC-18 column, 3 μm, 4.6 × 150 mm) using a linear gradient (0−20% or 0−30% for 7-mer and 0−30% for 23-mer) of buffer B during 40 min (A, 50 mM triethylammonium acetate in water; B, 50 mM triethylammonium acetate in MeCN−water, 1:1) and a flow rate of 1 mL/min. The anion-exchange purified oligonucleotides were dissolved in 0.5 mL of buffer A and filtered through a disposable syringe filter unit (Millex-GV13, 0.22 μm, 13 mm) before injection. The oligonucleotides were collected, lyophilized, redissolved in water, and lyophilized again.

The reagents and solvents used in synthesis were as follows: The building blocks [5′-*O*-(4-monomethoxytrityl)-thymidine 3′-H-phosphonate, 5′-*O*-(4-monomethoxytrityl)-2′-*O*-(*tert*-butyldimethylsilyl)uridine 3′-H-phosphonate, 5′-*O*-(4-monomethoxytrityl)-N^4-propionyl-2′-*O*-(*tert*-butyldimethylsilyl)cytidine 3′-H-phosphonate, 5′-*O*-(4-monomethoxytrityl)-N^6-butyryl-2′-*O*-(*tert*-butyldimethylsilyl)adenosine 3′-H-phosphonate, and 5′-*O*-(4-monomethoxytrityl)-N^2-(phenoxyacetyl)-2′-*O*-(*tert*-butyldimethylsilyl)guanosine 3′-H-phosphonate] were synthesized from the protected nucleosides using the PCl_3−imidazole−triethylamine system as described for deoxyribo- (Garegg et al., 1986a,c) and ribonucleosides (Garegg et al., 1986b; Stawinski et al., 1988) except that coevaporation with pyridine−triethylamine (4:1) before chromatography was omitted.

Pyridine (Lab Scan, Anhydro Scan) was stored over molecular sieves (4 Å). Acetonitrile and dichloroethane were of analytical quality (Merck) and stored over molecular sieves (3 and 4 Å, respectively). Pivaloyl chloride was distilled at atmospheric pressure and stored at −20 °C in a sealed flask. Trifluoroacetic acid was distilled at atmospheric pressure and stored at room temperature in a sealed flask. Triethylamine trihydrofluoride (TEA·3HF) was from Aldrich.

Both anion-exchange and reversed-phase HPLC were performed with Gilson equipment (units 305, 306, 805, 811C) and were detected by measuring UV absorbance at 260 nm with the same company's 117 UV detector. The water used when handling oligonucleotides was doubly distilled in a glass apparatus (acid washed every 6 months).

Oligodeoxyribonucleotides were obtained from Symbicom, Umeå, Sweden. They were synthesized according to the phosphotriester method, purified by HPLC, and freeze-dried from water.

Sample Preparation for Optical and Circular Dichroism Studies. Oligonucleotide solutions were prepared with a sterilized buffer system containing either 10 mM sodium phosphate and 0.1 mM Na_2EDTA (pH = 7.0) or 10 mM sodium cacodylate and 0.1 mM Na_2EDTA (pH = 7.0). Stock oligomer solutions were prepared by directly dissolving dry and desalted oligomers in the appropriate buffers. For each solution, the concentration in single strands was spectroscopically determined. The single strand absorbance at 25 °C was determined by extrapolation of the upper baseline in the UV melting curves. The corresponding 25 °C extinction coefficient was calculated by nearest neighbor analysis (Warshaw & Cantor, 1970; Puglisi & Tinoco, 1989). The concentrations of the oligomers were determined using the following extinction coefficients in single strands at 260 nm and 25 °C: 23-mer RNA (2.28×10^5 M^{-1} cm^{-1}); 23-mer RNA variant (2.31×10^5 M^{-1} cm^{-1}); 7-mer RNA, 7-mer RNA variant 1, and 7-mer RNA variant 2 (0.77×10^5 M^{-1} cm^{-1}); 23-mer DNA (2.21×10^5 M^{-1} cm^{-1}); 7-mer DNA (0.74×10^5 M^{-1} cm^{-1}). All concentrations are designated on a per strand basis unless otherwise mentioned.

UV Spectroscopy. Absorbance versus temperature profiles were obtained at 260 nm with a Cary 3E UV spectrophotometer. Quartz cells with path lengths of 1 and 0.2 cm were

used to allow measurements over a range of oligomer concentrations. The temperature of the sample was increased continuously at a rate of 0.5 °C per minute using an ethylene glycol water circulating bath (Lauda RM6) fitted with a thermoregulator (Eurotherm 808).

Thermodynamic Analysis. The UV melting curves were analyzed to obtain van't Hoff transition enthalpies. This analysis requires converting the experimental absorbance versus temperature curves into an *f* vs temperature curve, where *f* represents the fraction in single strands. The melted fraction (*f*) vs temperature (*t*) plots were obtained by fitting the melting profile to a two-state transition model, with linear sloping lower and upper baselines (Marky & Breslauer, 1987). The t_m's were obtained directly from the temperature at $f = 0.5$. A correction was done to take into account the difference between the bath and the sample temperature. Two methods were used to determine the transition enthalpy.

(a) From the van't Hoff Plot or ln K vs 1/ T Plot. Values of the equilibrium constant *K* were determined at each temperature using the appropriate equation. For a monomolecular transition this is simply $K = (1 - f)/f$ (Marky & Breslauer, 1987; Puglisi & Tinoco, 1989). Only points with $0.15 \leq f \leq 0.85$ were used in the van't Hoff plot since this is a region where the *K* values are most precise. For a two-state transition $\Delta H°$ is independent of temperature. The van't Hoff plot is linear with $(-\Delta H°/R)$ as the slope and $(\Delta S°/R)$ as the intercept. The slopes and the intercepts of the van't Hoff plots were obtained by linear regression analysis with at least 99% confidence level.

(b) From the Slope of the f vs T Plot at $T = T_m$. The $\Delta H°$ values were also obtained from the slope of the *f* vs *T* (K) melting curve at T_m using the relation (Marky & Breslauer, 1987)

$$\Delta H° = -(2 + 2n)RT_m^2(\delta f/\delta T)_{T=T_m}$$

where *n* is the molecularity of the association reaction and *R* is the gas constant. This relation is true for both self-complementary and non-self-complementary sequences.

Circular Dichroism Spectroscopy. CD spectra were recorded on a Jasco Model 720 spectropolarimeter equipped with a thermoelectrically controlled cell holder using the same samples as in the UV melting studies. Each spectrum reported here is an average of at least five scans. The samples were left for at least 10 min at 4 °C after each addition of salt before the CD spectra were recorded. The experiments probing interactions between 23-mer RNA and 7-mer RNA as well as those between the variants were repeated several times with different stock solutions of synthesized oligonucleotides.

Molecular Modeling Protocol. Modeling was performed using the graphics package INSIGHT II (Biosym Technologies, San Diego, CA) and also MIDAS PLUS (Computer Graphics Laboratory UCSF, San Francisco, CA). To model our system, we first prepared a RNA hairpin with a 3dT loop and the stem in the standard A-form, comprising the two helices P4 and P6 fused together to make them coaxial (Murphy et al., 1994). We prepared this hairpin following the protocol of Harvey et al. (1988). We then separated the part of the 7-mer RNA involved in the P6 helix from the rest of the hairpin, which now consists of the 23-mer RNA excepting the J3/4 portion. This is done by breaking the covalent connectivity at the position between U452 and U46 in Figure 1c. To facilitate the placement of the J6/7 part of the 7-mer and the J3/4 part of the 23-mer in the major and the minor groove of the the P4 and P6 helices, respectively, the P6 helix is rotated with respect to the P4 helix, keeping the part of the 23-mer RNA involved in the helices intact. This results in the alignment of the portion of the 7-mer RNA involved in the P6 helix along the major groove of the P4 helix, and automatically the minor groove of the P6 helix is aligned to accommodate the J3/4 part of the 23-mer RNA. Rotation between the P4 and P6 helices were also inferred by Chastain and Tinoco from their NMR studies (Chastain & Tinoco, 1992a, 1993). The individual bases of the J6/7 portion of the 7-mer RNA were then docked manually in the major groove of the P4 helix, keeping the base sequence and the covalent connectivities in the 7-mer RNA intact. Similarly, the bases of the J3/4 region of the 23-merRNA were also docked into the minor groove of the P6 helix. In the docking process we followed the triplet scheme as proposed in previous studies (Michel et al., 1990; Michel & Westhof, 1990) from comparative sequence analysis.

The initial structure was energy minimized for 500 Powell steps and was used as a starting structure for subsequent MD refinement. During MD simulation, the energy-minimized model system was first heated to 300 K over 1 ps, followed by equilibration at 300 K for 2 ps. A free MD simulation was then performed for 15 ps. The average structure over the last 5 ps was calculated and was energy minimized for 500 Powell steps. This structure was then used to analyze its qualitative structural features. All calculations were done in vacuum, using the CHARMM 22 package with nucleic acid parameters (Brooks et al., 1983). A distance-dependent dielectric constant and a reduced phosphate charge on the backbone were used to partially mimic the effects of solvents and counterions (Nilsson & Karplus, 1986).

RESULTS

UV Melting Curves. To investigate whether the 23-mer RNA and the 23-mer DNA form stable hairpin structures, we used the temperature dependence of the absorbance at 260 nm, as well as the salt dependence of the melting profiles, to characterize the type of transition involved. The helix/coil transition of the ordered structures formed by the 23-mer RNA, as well as the 23-mer DNA having the corresponding nucleotide sequence, was characterized by UV melting curves. The melting occurred in broad monophasic transitions covering a temperature range of 30 °C, with hypochromicities at 260 nm of 13% for 23-mer RNA and 12% for 23-mer DNA. Typical *f* vs temperature curves are presented in Figure 2. The transition temperature t_m (°C) did not change over a concentration range of 0.9−20.4 μM for 23-mer RNA and 0.8−43.3 μM for 23-mer DNA. This is consistent with the monomolecular melting of single-stranded intramolecular hairpins. The sequence (Figure 1b) suggests that seven stable hydrogen-bonded base pairs should contribute to the stem of the hairpin. The t_m of 23-mer RNA (50.9 °C) is higher than that of the corresponding 23-mer DNA (43.2 °C). Melting of the 7-mer RNA as well as 7-mer DNA showed no monophasic transitions, but the absorbance increased monotonically with temperature, with a total hypochromicity of 3−4%, which is also true for the variants of 7-mer RNA. This is characteristic of unstacking of single-stranded oligonucleotides.

Salt Dependence of t_m. Figure 3 shows the dependence of the transition temperature on salt concentration, in the

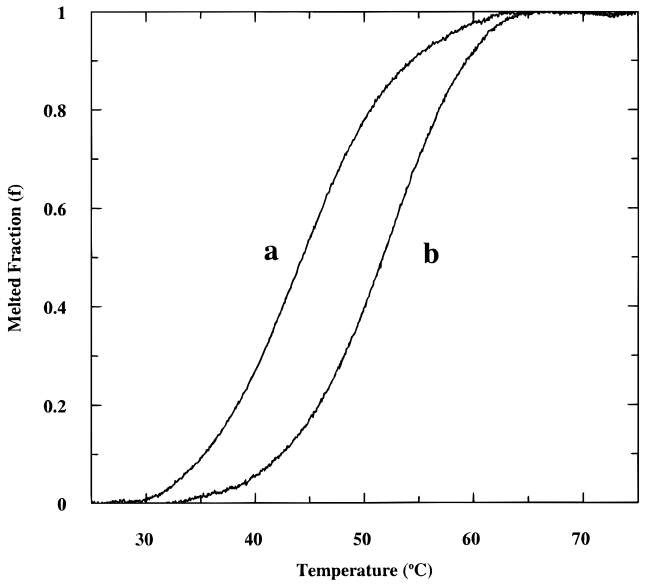

FIGURE 2: Typical melted fraction (f) vs temperature (t) curves of 23-mer DNA (a) and 23-mer RNA (b) in 10 mM sodium phosphate buffer, containing 0.1 mM Na_2EDTA at pH = 7.0. The oligomer concentration is 4.1 μM.

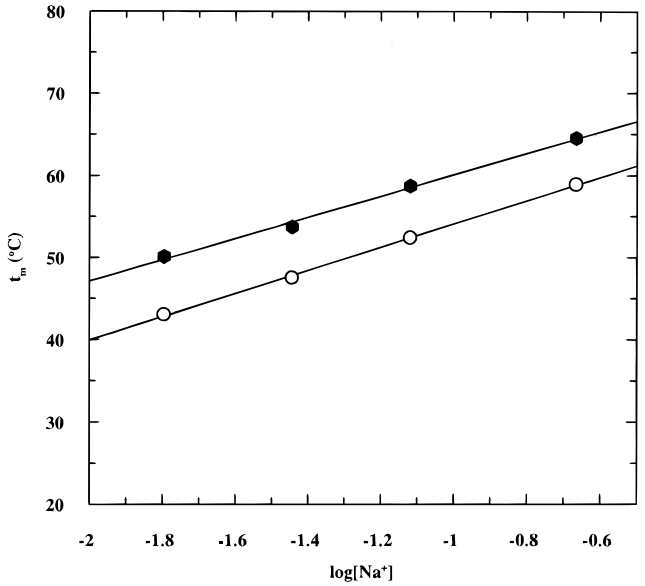

FIGURE 3: Salt dependence of the t_m for the 23-mer RNA hairpin (●) and 23-mer DNA hairpin (○) in 10 mM sodium phosphate buffer, containing 0.1 mM Na_2EDTA at pH = 7.0 adjusted to the desired NaCl concentration.

form of t_m vs log [Na^+] plots for 23-mer RNA and 23-mer DNA. The slopes of straight lines fitted to such curves are proportional to the differential number Δn of bound counterions between the single-stranded and duplex states (Record et al., 1978; Manning, 1978). The slopes for the t_m vs log [Na^+] are 13.0 °C for 23-mer RNA and 14.1 °C for 23-mer DNA, evaluated by linear regression. These values are characteristic of ion release associated with the melting of short oligomeric hairpins between 8 and 10 base pairs long (Rentzeperis et al., 1991; Zieba et al., 1991; Erie et al., 1987).

An estimate of Δn_{Na^+} per phosphate was made using the equation

$$dT_m/d \log [Na^+] = -0.9(2.303\, RT_m^2/\Delta H°)\Delta n$$

where the negative sign indicates a release of counterions. The value of 0.9 is a correction factor that corresponds to conversion of mean ionic activities to concentration, R is the universal gas constant in cal K^{-1} mol^{-1}, T_m is the transition temperature expressed in K, and $\Delta H°$ is the transition enthalpy in cal mol^{-1} of duplex. Δn is the counterion release on melting of one cooperative unit, which is equal to the dissociation of the entire duplex. The counterion released per phosphate is obtained by dividing Δn by the total number of phosphates involved in each transition (stem and loop for a hairpin) (Record, 1978). The $\Delta H°$ values used for these calculations are 48.3 kcal/mol for 23-mer RNA and 44.9 kcal/mol for 23-mer DNA. We obtained Δn_{Na^+} values per phosphate of 0.066 for 23-mer RNA and 0.070 for 23-mer DNA when we took into account all the phosphates of each molecule. Alternatively, if only the 14 helical phosphates of the hairpin stem are taken into account, the resulting Δn_{Na^+} values are 0.104 for 23-mer RNA and 0.110 for 23-mer DNA. These values are in good agreement with that obtained for DNA duplexes with eight base pairs ($\Delta n_{Na^+} = 0.086-0.131$) (Erie et al., 1987). The similar values of Δn_{Na^+} obtained for 23-mer RNA and 23-mer DNA indicate a similar type of hairpin formation in these two oligonucleotides. Since the loops are identical (three deoxythymidines) in the two hairpins, the small difference in the slopes and the Δn_{Na^+} values for the two hairpins should reflect a minor difference in the helical stem conformations.

Thermodynamic Profiles of Hairpin Formation. Table 1 lists the standard thermodynamic parameters for the formation of hairpins in 23-mer RNA compared with that of 23-mer DNA. The thermodynamic parameters calculated by the two different methods are in good agreement, which indicates that baselines are properly subtracted. Linear van't Hoff plots (ln K vs $1/T$) were obtained in all cases, which is consistent with the assumption of a two-state model. The 23-mer RNA hairpin is more stable than the 23-mer DNA as reflected in the t_m values. The free energy of formation of the 23-mer RNA hairpin at 25 °C is ~1.3 kcal/mol (in strands) more favorable than that of the 23-mer DNA hairpin. The small but significant increase in stability reflects the additional heat of formation that is not quite compensated by the loss in observed entropy. However, for both types of hairpins, the $\Delta H°$ values are in the same range, which reflect similar types of base pairing and stacking in both hairpin stems.

Circular Dichroism Spectroscopy. Figure 4 shows the CD spectra of the 23-mer RNA hairpin and 23-mer DNA hairpin at 4 °C. The CD spectrum of the 23-mer RNA hairpin is consistent with an A-form helix with a large positive band of ~250−300 nm and a large but narrower negative band at 212 nm (Johnson, 1985; Gray et al., 1992). The CD spectrum of the 23-mer RNA hairpin (Figure 4) with increasing temperature shows a decrease in the positive band of ~250−300 nm and also of the negative band of ~212 nm, until at 80 °C the CD spectrum has a shape typical of non-base-pairing oligonucleotides (data not shown). The presence of isodichroic points in these temperature-dependent profiles indicates that the hairpin melting essentially follows a two-state transition. The CD spectrum of the 23-mer DNA (Figure 4), with a sequence corresponding to that of the 23-mer RNA, is consistent with a B-form DNA helix. It shows a large positive band at ~275 nm and a negative band at ~250 nm. Increasing the temperature results in changes in the CD spectrum characteristic of the melting of a B-form helix (data not shown). The CD spectra of the 7-mer RNA

Table 1: Thermodynamic Parameters for the Formation of Hairpins in 23-mer RNA and 23-mer DNA[a,b]

		from van't Hoff plot				from slope of f vs T plot[c] at T_m			
str concn (μM)	t_m (°C)	$\Delta H°$ (kcal/mol)	$\Delta S°$ (eu)	ΔG_{298K} (kcal/mol)	ΔG_{310K} (kcal/mol)	$\Delta H°$ (kcal/mol)	$\Delta S°$ (eu)	ΔG_{298K} (kcal/mol)	ΔG_{310K} (kcal/mol)
				23-mer RNA Hairpin					
4.1−9.1	50.9	−48.2	−148.2	−4.0	−2.3	−48.4	−149.4	−3.9	−2.1
				23-mer DNA Hairpin					
0.8−43.3	43.2	−45.8	−144.2	−2.8	−1.1	−44.1	−141.6	−2.5	−0.9

[a] All thermodynamic parameters refer to the formation of hairpins in 10 mM sodium phosphate buffer and 0.1 mM Na_2EDTA (pH = 7.0). The values given in parentheses are the uncertainities of $\Delta H°$ (±6%), $\Delta S°$ (±5%), and t_m (±0.6 °C). [b] All thermodynamic parameters refer to the average over the concentration range mentioned. [c] The melted fraction (f) vs temperature plot is obtained by fitting the UV melting curves to a modified two-state model with linear sloping upper and lower baselines.

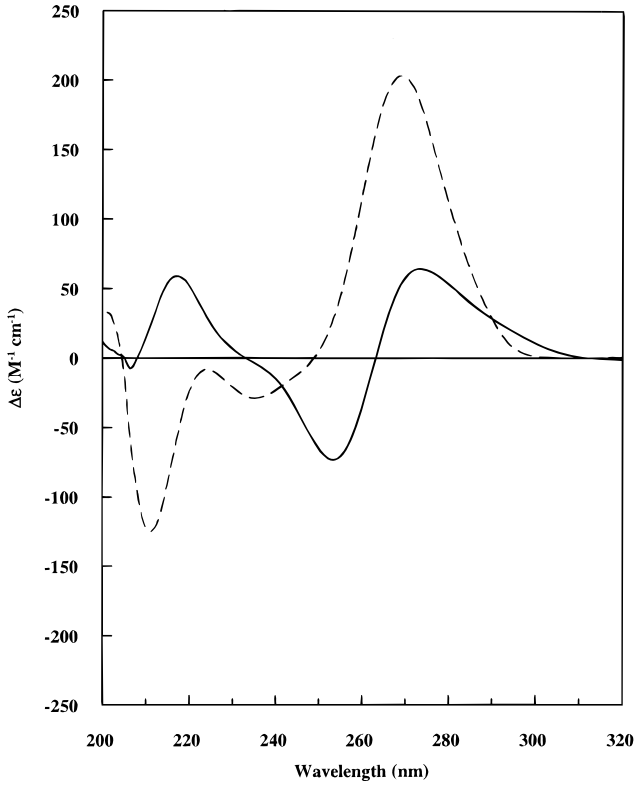

FIGURE 4: CD spectra of the 23-mer RNA hairpin (- - -) and 23-mer DNA hairpin (—) in 10 mM sodium phosphate buffer, containing 0.1 mM Na_2EDTA at pH = 7.0 at 4 °C. The strand concentration is ~5.0 μM.

and 7-mer DNA were characteristic of non-base-pairing oligonucleotides (Johnson, 1985).

Association between 23-mer RNA and 7-mer RNA. CD spectroscopy is very sensitive to the changes in nucleic acid backbone conformation. We therefore studied the association between 23-mer RNA and 7-mer RNA by observing changes in the CD spectrum, which result from mixing the two samples. The CD spectra for 23-mer RNA, 7-mer RNA, and a 1:1 mixture of them were recorded in different salt conditions to identify possible conditions for complex formation. For all the salt concentrations studied, it was ensured that the 23-mer RNA, 23-mer DNA, and the 23-mer RNA variant formed stable hairpins under the conditions used. The observed spectra of the 1:1 mixture of the 23-mer RNA hairpin and 7-mer RNA single strand were compared with the calculated spectra, obtained by adding the CD spectra of equimolar concentrations of 23-mer RNA hairpin and 7-mer RNA recorded separately under the same conditions as that of the 1:1 mixture in 10 mM sodium cacodylate buffer in the absence of Mg^{2+}. Care was taken to maintain identical experimental conditions to record the CD spectra of the 23-mer RNA, 7-mer RNA, and 1:1 mixture to ensure that no artifact was introduced for the comparison. When no association occurred, the comparison resulted in complete overlap of the observed and calculated spectra of the 1:1 mixture of the 23-mer RNA hairpin and the 7-mer RNA single strand. Figure 5a shows the overlayed observed and calculated spectra of the 1:1 mixture. It is evident that no association occurred between the 23-mer RNA and 7-mer RNA in the absence of Mg^{2+}. Our results are in good agreement with those of Chastain and Tinoco (1992b), in which complete overlap of calculated and observed CD spectra was obtained in the absence of triple helix formation between poly(rA) and poly(rG)·poly(rC). Similar results, indicative of absence of association, were also obtained for our system when monovalent cations were added (NaCl up to 500 mM).

Figure 5b shows the CD spectra of the 1:1 mixture of 23-mer and 7-mer RNA in 100 mM $MgCl_2$. In this case the observed spectrum is significantly different from the calculated spectrum of the complex, showing an increase in the positive peak at ~270 nm accompanied by a slight narrowing of the positive band at ~250−300 nm (Figure 5b). The negative band ~210 nm is also larger in case of the observed spectrum of the mixture. Our results agree well with those for triple helix formation between poly(rA) and poly(rG)·poly(rC) where similar qualitative differences between observed and calculated spectra of the 1:1 complex of poly(rA) and poly(rG)·poly(rC) were recorded (Chastain & Tinoco, 1992b). This difference between the observed and calculated spectrum of the 1:1 mixture in our system is indicative of an association between the 23-mer RNA hairpin and the 7-mer RNA, which may involve triplex formation with additional base-stacking interactions. The difference between the observed and calculated spectra at 269 nm has also been studied with an increasing fraction of 7-mer RNA, keeping the total strand concentration constant at 5.8 μM and in presence of 100 mM $MgCl_2$ (data not shown). The maximum value of the difference was observed at 0.5 fraction, indicating that the association stoichiometry is 1:1 and was complete under the conditions studied. In this connection it should be mentioned that the 1:1 mixture of 23-mer RNA and 7-mer RNA did not show any biphasic melting curve under all conditions studied.

Figure 5c shows the difference between the observed and calculated CD spectra of the 1:1 23-mer RNA−7-mer RNA mixture in different concentrations of $MgCl_2$. The difference spectrum in the absence of $MgCl_2$, when no association takes place, follows the baseline. Increasing the concentration of $MgCl_2$ results in enhancement of the positive peak around

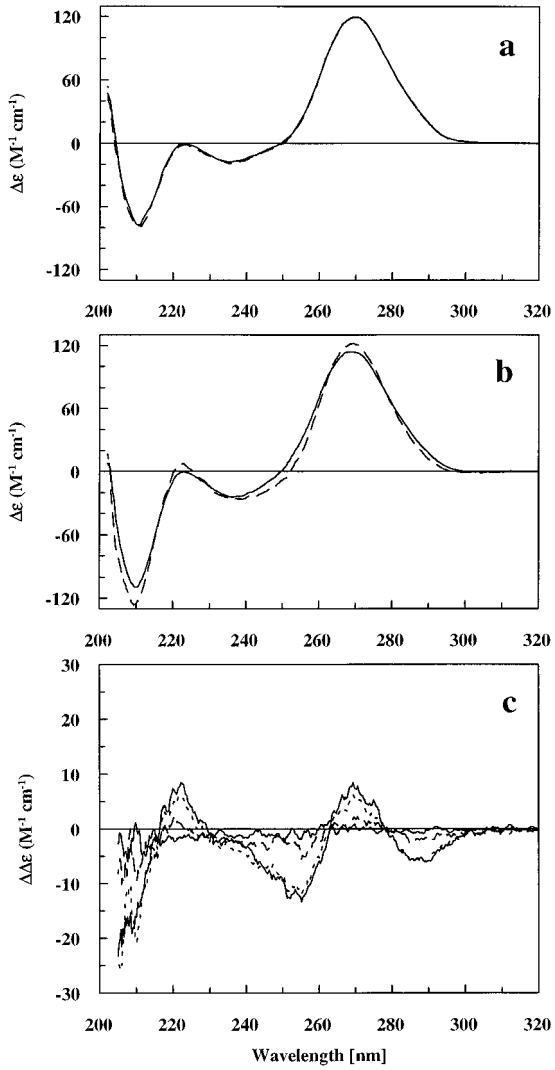

FIGURE 5: Overlayed CD spectra of a 1:1 mixture of 23-mer RNA and 7-mer RNA (- - -) and weighted sum of the spectra of 23-mer RNA and 7-mer RNA recorded separately under the same conditions (—) in the absence of added $MgCl_2$ (a) and in presence of 100 mM $MgCl_2$ (b). (c) Difference CD spectra of the 1:1 complex of 23-mer RNA and 7-mer RNA (observed − calculated) (see text) in 0 mM (—), 10 mM (- - -), 50 mM (•••), and 100 mM (—) $MgCl_2$. All solutions contain 10 mM sodium cacodylate and 0.1 mM Na_2-EDTA, pH = 7.0 at 4 °C. Oligomer strand concentrations were 5−6 μM.

270 nm and also the negative peak ~210 nm, which reflects an increased degree of association between the RNA oligonucleotides. We also observed that the association was completed on a time scale of less than 1 min under the present conditions (data not shown). Figure 6 shows the increase in differences between the observed and calculated CD spectra of 23- and 7-mer RNAs at 269 nm upon increasing the concentration of added $MgCl_2$. We observed that addition of $MgCl_2$ up to 200 mM led to a pH decrease from 7.0 to 6.6 in the samples. Since protonation of cytosines occurs upon lowering of pH, which might promote association of the RNA oligonucleotides, we investigated the effect of pH changes on our results. When the pH was gradually changed in a titration from 7.0 to 5.5 by adding HCl to a sample without Mg^{2+}, no evidence of association between the 7-mer and 23-mer RNA was seen (data not shown). Thus we concluded that the Mg^{2+}-induced association is not a pH effect. It should be mentioned that Mn^{2+} instead of Mg^{2+} could also induce association. The results were similar to that observed in case of Mg^{2+}. When a 1:1

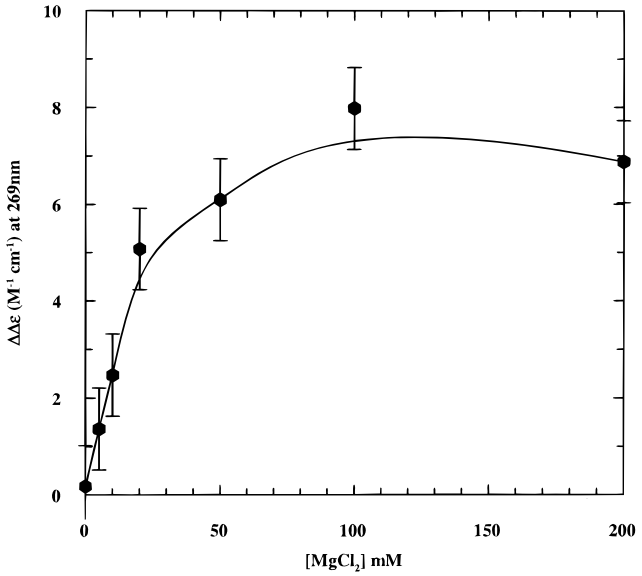

FIGURE 6: Change in $\Delta\epsilon$ values monitored at 269 nm, from difference CD spectra of the 1:1 complex of 23-mer RNA and 7-mer RNA (observed − calculated) (see text) upon increasing $MgCl_2$ concentration.

mixture of a 23-mer DNA hairpin and a 7-mer DNA single strand was studied under identical conditions, no significant difference was seen between the observed and calculated spectra of the 1:1 mixture at any concentration of $MgCl_2$ (0−100 mM). Hence we conclude that no association took place in this case.

Association between Variants of 23-mer RNA and 7-mer RNA. We have studied the association between 23-mer RNA and 7-mer RNA variants, not only to probe if our model system is sensitive to base substitutions to demonstrate the formation of the interactions proposed in our model of the triple-helical domain in Figure 1c but also to see if it is partly sensitive to sequence variations related to function as suggested from phylogenetic comparisons.

(i) 7-mer RNA Variant 1 and 23-mer RNA. In this 7-mer RNA variant the G451 in the original intron sequence (as shown in Figure 1c) is replaced by U in the 7-mer strand. This base substitution results in disruption of the proposed Watson−Crick base pair C-G in the P6 helix region of our model along with a P6−J3/4 nucleoside triple (Figure 1c) thus perturbing the formation of the P6 helix. Hence it is expected that this base substitution will weaken the association between the 23-mer RNA and 7-mer RNA variant 1. Figure 7a shows the overlayed CD spectrum of a 1:1 mixture of 23-mer RNA and 7-mer RNA variant 1 with that of the added spectra recorded separately. There is an almost complete overlap at 100 mM concentration of $MgCl_2$, indicating the lack of any association between the fragments. This is also reflected in the CD difference spectra recorded at increasing $MgCl_2$ concentration (Figure 7b).

(ii) 7-mer RNA Variant 2 and 23-mer RNA. Here the U453 in the original intron sequence which is on the 7-mer strand is replaced by G. This would disrupt a phylogenetically proposed major groove nucleoside triple interaction (Michel & Westhof, 1990) between the P4 helix and the J6/7 single-stranded region of our model (Figure 1c). The Watson−Crick base-paired region between the 23-mer and 7-mer RNA, i.e., the P6 part, remains undisturbed. Figure 8a (i) shows the overlayed observed and calculated spectra of a 1:1 mixture of 23-mer RNA and 7-mer RNA variant 2

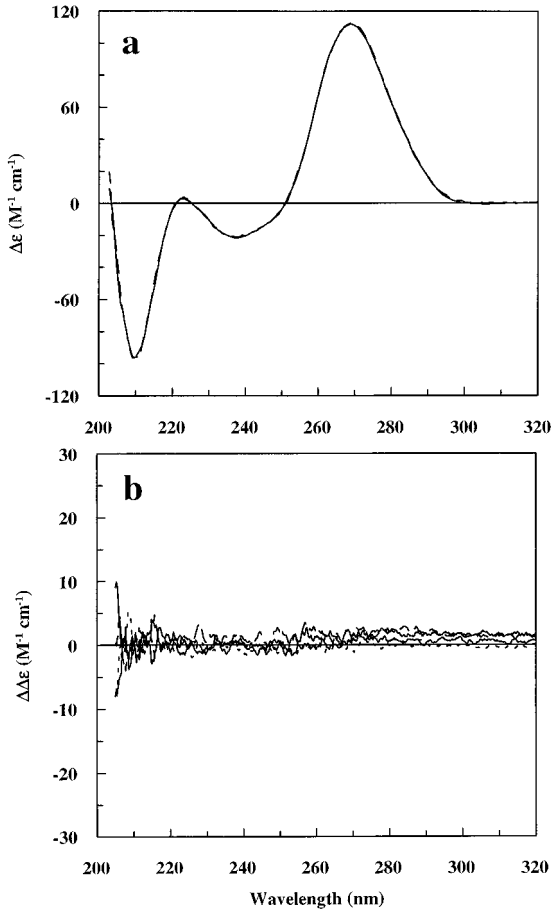

FIGURE 7: (a) Overlayed CD spectra of a 1:1 mixture of 23-mer RNA and 7-mer RNA variant 1 (- - -) and the weighted sum of the spectra of 23-mer RNA and 7-mer RNA variant 1 recorded separately under the same conditions (—) in 100 mM $MgCl_2$. (b) Difference CD spectra of the 1:1 complex of 23-mer RNA and 7-mer RNA variant 1 (observed − calculated) (see text) in 0 mM (—), 5 mM (- - -), 50 mM (•••), and 100 mM (−) $MgCl_2$. All solutions contain 10 mM sodium cacodylate and 0.1 mM Na_2EDTA, pH = 7.0 at 4 °C.

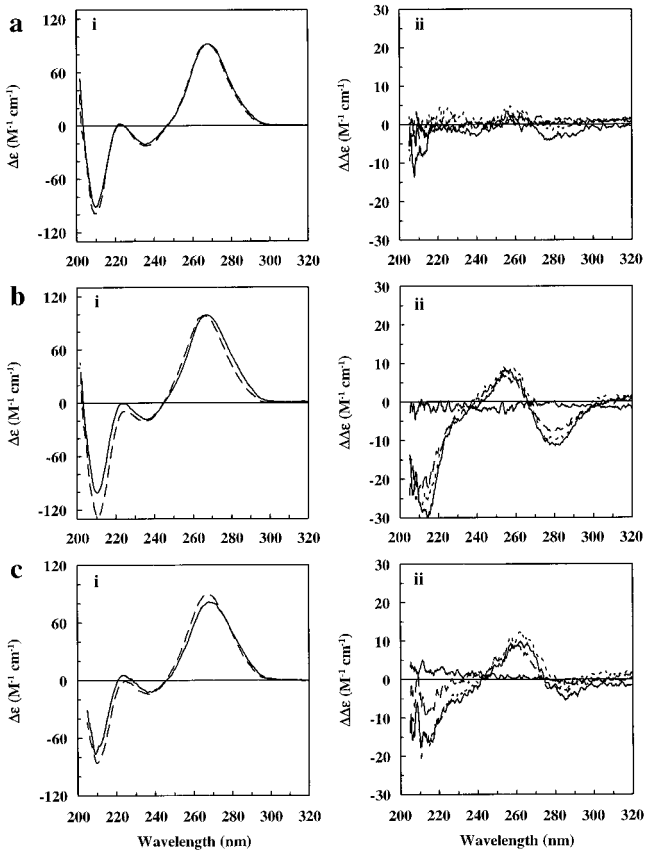

FIGURE 8: (a) (i) Overlayed CD spectra of a 1:1 mixture of 23-mer RNA and 7-mer RNA variant 2 (- - -) and the weighted sum of the spectra of 23-mer RNA and 7-mer RNA variant 2 recorded separately under the same conditions (—) in 100 mM $MgCl_2$. (ii) Difference CD spectra of the 1:1 complex of 23-mer RNA and 7-mer RNA variant 2 (observed − calculated) (see text) in 0 mM (—), 5 mM (- - -), 50 mM (•••), and 100 mM (−) $MgCl_2$. (b) (i) Overlayed CD spectra of a 1:1 mixture of 23-mer RNA variant and 7-mer RNA variant 2 (- - -) and the weighted sum of the spectra of 23-mer RNA variant and 7-mer RNA variant 2 recorded separately under the same conditions (—) in 100 mM $MgCl_2$. (ii) Difference CD spectra of the 1:1 complex of 23-mer RNA variant and 7-mer RNA variant 2 (observed − calculated) (see text) in 0 mM (—), 5 mM (- - -), 50 mM (•••), and 100 mM (−) $MgCl_2$. (c) (i) Overlayed CD spectra of a 1:1 mixture of 23-mer RNA variant and 7-mer RNA (- - -) and the weighted sum of the spectra of 23-mer RNA variant and 7-mer RNA recorded separately under the same conditions (—) in 100 mM $MgCl_2$. (ii) Difference CD spectra of the 1:1 complex of 23-mer RNA variant and 7-mer RNA (observed − calculated) (see text) in 0 mM (—), 5 mM (- - -), 50 mM (•••), and 100 mM (−) $MgCl_2$. All solutions contain 10 mM sodium cacodylate and 0.1 mM Na_2EDTA, pH = 7.0 at 4 °C. Oligomer strand concentrations ~4−6 μM.

in 100 mM $MgCl_2$. A very small peak shift (<1 nm) is seen in the observed spectrum compared with that of the calculated spectrum, along with a small increase in the negative peak at ~210 nm. The CD difference spectrum with increasing concentration of $MgCl_2$ [Figure 8a (ii)] shows that even in the presence of 50 mM $MgCl_2$ there is no association and the difference spectrum follows the baseline. The small difference seen at concentrations of 100 mM and above could be due to a very weak association between the 23-mer RNA and 7-mer RNA variant 2.

(iii) 23-mer RNA Variant and 7-mer RNA Variant 2. In the 23-mer RNA variant the G47 is replaced by C and C74 is replaced by G such that the second Watson−Crick base pair of the P4 helix region is now C-G instead of G-C. This double base substitution on 23-mer RNA is aimed at removing the disruptive effect of the base substitution done in 7-mer RNA variant 2, such that the association between the 23-mer RNA variant and 7-mer RNA variant 2 is favored, which would restore the P4−J6/7 nucleoside triple interaction (C·G·G), whose existence in the group 1 introns has been predicted by comparative sequence analysis (Michel & Westhof, 1990). Figure 8b (i) shows that the observed spectrum of the 1:1 mixture of the 23-mer RNA variant and 7-mer RNA variant 2 is significantly different from that of the calculated spectrum, indicating association between the two. There is a 2 nm shift in the positive peak along with a large increase in the negative peak at ~210 nm. Figure 8b (ii) shows the increase in the differences between the observed and calculated spectra with increasing concentration of $MgCl_2$. Association occurs even for concentrations as low as 5 mM $MgCl_2$. No further change in the difference spectrum occurs above 100 mM $MgCl_2$.

(iv) 23-mer RNA Variant and 7-mer RNA. Association was observed between the 23-mer RNA variant and 7-mer RNA [Figure 8c (i and ii)]. Comparison of the observed spectrum of the 1:1 mixture of the 23-mer RNA variant and 7-mer RNA with the calculated spectrum shows a small shift in the positive peak (~1 nm) along with an increase in the negative at ~210 nm [Figure 8c (i)], which is however less than that observed in the previous case [Figure 8b (i and ii)]. Figure 8c (ii) shows the increase in the differences

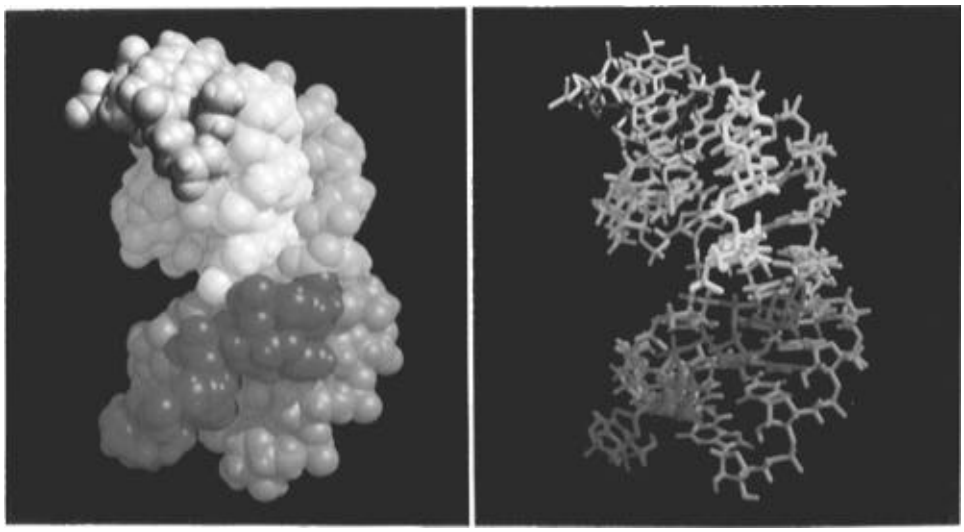

FIGURE 9: Space-filled (a, left) and stick (b, right) representation of the average energy-minimized structure of the model triple helical domain obtained from the last 5 ps of the MD simulation. The P4 helix is shown in blue, the P6 helix in green, the J3/4 in red, the J6/7 in yellow, and the 3dT loop in gray.

between the observed and calculated spectra with increasing $MgCl_2$ concentration. Our results indicate that changing the second base pair of the P4 helix from G-C to C-G in the 23-mer RNA has a less perturbing effect on the association than changing U453 to G in the 7-mer RNA.

It should also be mentioned that in the absence of magnesium no association was detected irrespective of all the variants studied and the CD difference spectra followed the baseline in all cases.

Molecular Modeling. Figure 1c shows a schematic representation of the P4−P6 along with the J3/4 and J6/7 regions that our physical model system might adopt. It is based on previously suggested models (Michel & Westhof, 1990; Chastain & Tinoco, 1992a, 1993). The present molecular modeling includes 15 ps MD simulation and energy minimization following the protocol described in the Materials and Methods section. The results show (Figure 9) that it is possible to form a structure accommodating the J3/4 region of the 23-mer RNA in the minor groove of the P6 helix and the J6/7 region of the 7-mer RNA in the major groove of the P4 helix without any steric clashes. It was necessary to rotate the P6 helix with respect to P4, in order to accommodate the J3/4 and the J6/7 in the minor and the major grooves of the P6 and P4 helices, respectively.

In the final energy-minimized structure (Figure 9), both the P4 and P6 helices remain intact except for the G·U base pair of the P6 and the junction base pair (U·A) of the P4 helix (Figure 1c), which are broken. These two base pairs are involved in the crossover region of our structure (Figure 9). In the crossover region after MD simulation and energy minimzation, there appeared some unusual H-bonds which involve the 2′-OH groups of the ribose sugars of the junction nucleotides. In our final energy-minimized average structure we have also found that most of the bases of the J3/4 and J6/7 regions interact with the P6 and P4 helices through H-bonds. We have also found that the 2′-OH groups play an important role, particularly in interactions between J3/4 and the minor groove of P6 helix and to some extent also in interactions between J6/7 and the major groove of the P4 helix. In addition, the minor groove of P6 and the major groove of P4 are found to be slightly widened to accommodate the respective single-stranded regions. A similar modeling study on the corresponding DNA system showed that the J3/4 segment did not interact at all with the minor groove of the P6 helix. The role of the individual hydroxyl groups in the tertiary interactions of the RNA model system will be the subject of further experimental and modeling studies.

DISCUSSION

Hairpin Formation in 23-mer RNA and 23-mer DNA. The sequences of 23-mer RNA and 23-mer DNA (Figure 1) favor the formation of intramolecular hairpins involving seven base pairs interspaced with an internal loop of three deoxythymidines. The t_m's of helix to coil transition remained unaffected by increases in the strand concentration (from 0.9 to 20.4 μM in case of 23-mer RNA and from 0.8 to 43.3 μM for 23-mer DNA), confirming that the structures formed at low temperature are monomolecular. A monomolecular structure is also consistent with the much higher t_m of 70.5 °C in 1 M NaCl for 23-mer RNA hairpin (data not shown) that is observed relative to the estimated t_m of 23 °C based on the nearest neighbor interactions (Turner et al., 1988) for the stem duplex (1 μM in strand concentration) in this salt concentration. In general, the t_m of hairpin melting is higher than the t_m of the stem duplex melting (Rentzeperis et al., 1991, 1993). The transition enthalpy of ∼−48 kcal/mol for 23-mer RNA and ∼−45 kcal/mol for 23-mer DNA is in good agreement with the estimated transition enthalpy of the hairpin stem (∼−46 kcal/mol), calculated by assuming that both 23-mer RNA and 23-mer DNA form hairpins with three deoxythymidines in the loop and G-C as the loop closing base pair. The salt dependence of the t_m is also characteristic for hairpins with short stems of about 8−10 base pairs (Rentzeperis et al., 1991; Zieba et al., 1991; Erie et al., 1987). The observed Δn_{Na^+} values in the 23-mer RNA and 23-mer DNA provide additional evidence for the formation of intramolecular hairpins with short stems.

The 23-mer RNA and 23-mer DNA have equivalent sequences, and it is therefore expected that they would form hairpins which are very similar, having the same loop sequence (3dT), loop closing base pair (G-C), and the corresponding sequence in the stem. The close values of transition enthalpies, the dependence of t_m on salt concentration, and the counterion release all strongly point to this fact.

However, there is a difference in the thermodynamic stability between the two hairpins. This is reflected in the higher t_m value of the 23-mer RNA compared to that of 23-mer DNA (Figure 2; Table 1). The stability of a RNA/DNA hairpin is mainly guided by the loop sequence and the closing base pair (Turner et al., 1988; Sera et al., 1993). In the present case, both hairpins have the same loop and the same bases involved as the closing base pair. Therefore, the difference in the thermodynamic stability should be accounted for by the difference in the stem duplexes.

From the CD studies (Figure 4) it is evident that the stems of the two hairpins adopt quite different conformations (A helix for 23-mer RNA and B helix for 23-mer DNA). This could account for the difference in thermodynamic stability and also for the differences in the other parameters determined.

Triplex Formation between 23-mer RNA and 7-mer RNA. In the absence of any association between the 23-mer RNA hairpin and 7-mer RNA, a complete overlap was seen between the observed spectrum of the 1:1 mixture and the spectrum calculated as the sum of the separate components (Figure 5a). No association was detected in the presence of monovalent cations (up to 500 mM NaCl). However, association occurred and could be detected above 5 mM $MgCl_2$ (Figure 5c). The qualitative difference between observed and calculated spectra shows a similarity to that observed for triple helix formation between poly(rA) and poly(rG)·poly(rC) (Chastain and Tinoco, 1992b). In experiments where both or one of the participating molecules was a deoxyribonucleotide no evidence of association was found. This would imply that the presence of the 2′-OH group in the ribose sugar might be partially involved in hydrogen bonding to form the required nucleoside triples for the association [cf. Whoriskey et al. (1995) and Chastain and Tinoco (1992a, 1993)]. Further work to probe this aspect is in progress and will be communicated separately.

The proposed structure of the present model system is shown in Figure 1c and illustrated in Figure 9. In the proposed structure, three out of the first four bases from the 5′-end of the 7-mer form Watson−Crick base pairs with the 3′-end bases of the 23-mer RNA to model the P6 helix region. The six base-paired stem of the 23-mer RNA hairpin models the P4 helix. The loose ends of J3/4 (5′-end of 23-mer RNA) and J6/7 (3′-end of 7-mer RNA) associate with P6 and P4 helices, respectively, with the nucleoside triple interactions as shown in Figure 1c. The nucleoside triple interactions in our model system are also suggested by phylogenetic and mutagenic evidence (Michel & Westhof, 1990; Green & Szostak, 1994). The model is also in agreement with the results of an NMR study of a corresponding 25-mer oligoribonucleotide that models the junction of coaxially stacked P4/P6 helices with the J3/4 sequence associated as a triplet-forming strand in the minor grove of the P6 helix (Chastain & Tinoco, 1992a, 1993).

Studies on variants of 7-mer and 23-mer RNA were carried out to investigate the response of the model to base substitutions, by changing some proposed interactions to elucidate their importance in this kind of triple strand association. The two 7-mer RNA variants were designed to perturb the association at two different positions. For the 7-mer RNA variant 1 the single base replacement (G451 by U, Figure 1c) should result in disruption of the Watson−Crick base pair in the P6 helix of our model, perturbing the association between 23-mer RNA and 7-mer RNA variant 1. In order to show the importance of nucleoside triple interaction for the association of our two strands, we then studied the 7-mer RNA variant 2, where changing U453 to G results in a nucleoside triple whose existence is not found among the phylogenetic variations. It is expected to disrupt one nucleoside triple interaction between P4−J6/7 but leave the P6−J3/4 region and P4 helix of the model system unperturbed. Hence weakening of the association between 23-mer RNA and the 7-mer RNA variant 2 could be expected. The double base substitution on 23-mer RNA (23 mer RNA variant 1) that changes the second base pair in the P4 helix from G-C to C-G was aimed at removing the perturbing effect of the 7-mer RNA variant 2 by restoring the P4−J6/7 nucleoside triple to C·G·G, whose existence in group 1 introns was predicted by comparative sequence analysis (Michel & Westhof, 1990). This multiple base substitution would restore the association between the 23-mer RNA variant and 7-mer RNA variant 2. Association between the 23-mer RNA variant and 7-mer RNA was carried out to probe the sensitivity of our model to the change in the base pair (G-C to C-G) on the P4 helix alone and was aimed at disrupting a phylogenetically predicted nucleoside triple interaction.

Our CD results on the model system show that true to our expectations (a) there is no association when the formation of the P6 helix is perturbed by a single base replacement of G451 to U (Figure 7a,b), (b) only a very weak association in high $MgCl_2$ ($\geq$100 mM) was observed when the phylogenetically proposed nucleoside triple interaction was disrupted by changing U453 to G [Figure 8a (i and ii)] on the 7-mer RNA (this shows that, for the association between our two strands, the nucleoside triple interactions play an important role), and (c) association was restored even in low $MgCl_2$ ($\sim$5 mM) when the double base substitution on the 23-mer RNA (G-C to C-G) revived the nucleoside triple interaction to C·G·G, whose existence has been proposed from comparative sequence analysis (Michel & Westhof, 1990; Green & Szostak, 1994). However, although not predicted by phylogenetic data, association does occur in the model system when only the second base pair G-C was changed to C-G in the P4 helix [Figure 8c (i and ii)]. The constraints that determine the identity of the second and third base pair of the P4 helix are likely to be located outside the triple helical domain as pointed out by previous workers (Green & Szostak, 1994). This could explain why inversion of the second base pair of the P4 helix region in our model does not abolish the interaction although it is not phylogenetically predicted. However, the fact that our base substitution results to a certain degree follow a general pattern predicted by phylogenetic results demonstrates the sensitivity of this model system in studying the effect of mutations on these kinds of tertiary contacts in the group 1 intron core. These results show that our two-strand system with added Mg^{2+} may serve as a good structural model for the triple helical domain in the catalytic core of the group 1 intron. We have also shown the Mg^{2+} dependence of the association. The triplet interactions of the model in Figure 1c are required for the proper functions in the splicing reactions of a group I intron (Green & Szostak, 1994), which also requires the presence of Mg^{2+} for the splicing activity and tertiary structure stabilization (Murphy et al., 1994). Mg^{2+}-induced association as observed in this case is of particular interest in the light of the substitution interference experiments with nucleoside phosphorothioates on the group 1 intron from *T.*

thermophila (Christian & Yarus, 1993). These results identified a number of magnesium binding sites that significantly influenced the rate of the first step of splicing (G-addition). Most intriguing and perhaps a direct correlation between structure and function of the ribozyme is that three of these Mg^{2+} binding sites are located within the part of the catalytic core that correspond to our 7-mer RNA model (nucleotides 449−455). It should be mentioned that the G-binding site is located near the triple helical domain at G458 of P7.

Our molecular modeling shows that the entire proposed structure (Figure 1c) is energetically possible without any steric clashes in which most of the bases of the J3/4 and J6/7 regions can interact with the minor groove of P6 and major groove of P4 helices, respectively, giving a possible novel structural element based on a pseudoknot and resembling a plaited triple helical structure (Figure 9). The experiments on deoxyribonucleotides and the molecular modeling indicate that the 2′-OH groups of the ribose sugar play an important role in forming the structure.

CONCLUSIONS

A new aspect of our present model system, for the P4−P6 triple helical domain of the group I intron core, is that it involves association of *two separate RNA strands* based on native intron sequences with specific sequence requirements to guide the tertiary interactions. Mg^{2+} is required for the association, and this correlates well with the reported specific and functionally important Mg^{2+} binding sites localized in the J6/7 region of the intron (Christian & Yarus, 1993). The studies on different sequence variants show that, for the interaction to occur between the two strands in the model system, not only is the formation of the Watson−Crick base pairs in the P6 and P4 helices essential, but also nucleoside triple interactions are required. The base substitution results seem to follow a general pattern predicted by comparative sequence analysis of group 1 introns. Our model system is therefore a good synthetic model for further studies of the structural constraints that exist in this plaited triple helical domain of a functional RNA molecule.

ACKNOWLEDGMENT

We thank Dr. L. Nilsson for helpful discussions and Mr. E. Westman for technical assistance.

REFERENCES

Atkinson, T., & Smith, M. (1984) in *Oligonucleotide Synthesis—A Practical Approach* (Gait, M. J., Ed.) pp 35−81, IRL Press, Oxford.

Baxter, S. M., Greizerstein, M. B., Kushlan, D. M., & Ashley, G. W. (1993) *Biochemistry 32*, 8702−8711.

Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S., & Karplus, M. (1983) *J. Comput. Chem. 4*, 187−217.

Cech, T. R., Damberger, S. H., & Gutell, R. R. (1994) *Nat. Struct. Biol. 1*, 273−280.

Celander, D. W., & Cech, T. R. (1991) *Science 251*, 401−407.

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

Chastain, M., & Tinoco, I., Jr. (1992b) *Nucleic Acids. Res. 20*, 315−318.

Chastain, M., & Tinoco, I., Jr. (1993) *Biochemistry 32*, 14220−14228.

Christian, E. L., & Yarus, M. (1993) *Biochemistry 32*, 4475−4480.

Couture, S., Ellington, A. D., Gerber, A. S., Cherry, J. M., Doudna, J. A., Green, R., Hanna, M., Pace, U., Rajagopal, J., & Szostak, J. W. (1990) *J. Mol. Biol. 215*, 345−358.

Davies, R. W., Waring, R. B., Ray, J. A., Brown, T. A., & Scazzocchio, C. (1982) *Nature 300*, 719−724.

Doudna, J. A., & Cech, T. R. (1995) *RNA 1*, 36−45.

Erie, D., Sinha, N., Olson, W., Jones, R., & Breslauer, K. J. (1987) *Biochemistry 26*, 7150 −7159.

Garegg, P. J., Lindh, I., Regberg, T., Stawinski, J., Strömberg, R., & Henrichson, C. (1986a) *Tetrahedron Lett. 27*, 4051−4054.

Garegg, P. J., Lindh, I., Regberg, T., Stawinski, J., Strömberg, R., & Henrichson, C. (1986b) *Tetrahedron Lett. 27*, 4055−4058.

Garegg, P. J., Regberg, T., Stawinski, J., & Strömberg, R. (1986c) *Chem. Scr. 26*, 59−62.

Gasparutto, D., Livache, T., Bazin, H., Duplaa, A.-M., Guy, A., Khorlin, A., Molko, D., Roget, A., & Teoule, R. (1992) *Nucleic Acids Res. 20*, 5159−5166.

Gray, D. M., Ratliff, R. L., & Vaughan, M. R. (1992) *Methods Enzymol. 211*, 389−405.

Green, R., & Szostak, J. W. (1994) *J. Mol. Biol. 235*, 140−155.

Harvey, S., Luo, J., & Lavery, R. (1988) *Nucleic Acids Res. 16*, 11795−11809.

Inoue, T., & Cech, T. R. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 648−652.

Johnson, W. C., Jr. (1985) *Methods Biochem. Anal. 31*, 61−163.

Manning, G. S. (1978) *Q. Rev. Biophys. 11*, 179−246.

Marky, L. A., & Breslauer, K. J. (1987) *Biopolymers 26*, 1601−1620.

Michel, F., & Westhof, E. (1990) *J. Mol. Biol. 216*, 585−610.

Michel, F., Jaquier, A., & Dujon, B. (1982) *Biochimie 64*, 867−881.

Michel, F., Hanna, M., Green, R., Bartel, D. P., & Szostak, J. W. (1989) *Nature 342*, 391−395.

Michel, F., Ellington, A. D., Courture, S., & Szostak, J. W. (1990) *Nature 347*, 578−580.

Murphy, F. L., Wang, Y., Griffith, J. D., & Cech, T. R. (1994) *Science 265*, 1709−1712.

Nilsson, L., & Karplus, M. (1986) *J. Comput. Chem. 7*, 591.

Pleij, C. W. A., Rietveld, K., & Bosch, L. (1985) *Nucleic Acids Res. 13*, 1717−1731.

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

Pyle, A. M., Murphy, F. L., & Cech, T. R. (1992) *Nature 358*, 123−128.

Record, M. T., Jr., Anderson, C. F., & Lohman, T. M. (1978) *Q. Rev. Biophys. 11*, 103−178.

Rentzeperis, D., Kharakoz, D. P., & Marky, L. A. (1991) *Biochemistry 30*, 6276−6283.

Rentzeperis, D., Alessi, K., & Marky, L. A. (1993) *Nucleic Acids Res. 21*, 2683−2689.

Rozners, E., Westman, E., & Strömberg, R. (1994) *Nucleic Acids Res. 22*, 94−99.

Rozners, E., Sigurdsson, S., Bizdena, E., Westman, E., & Strömberg, R. (1995) *Nucleosides Nucleotides 14*, 875−878.

Serra, M. J., Lyttle, M. H., Axenson, T. J., Schadt, C. A., & Turner, D. H. (1993) *Nucleic Acids Res. 21*, 3845−3849.

Sjöberg, B. M., Hahne, S., Mathews, C. Z., Mathews, C. K., Rand, K. N.. & Gait, M. J. (1986) *EMBO J. 5*, 2031−2036.

Sproat, B. J. (1984) in *Oligonucleotide Synthesis—A Practical Approach* (Gait, M. J., Ed.) pp 83−115, IRL Press, Oxford.

Stawinski, J., Strömberg, R., Thelin, M., & Westman, E. (1988) *Nucleic Acids Res. 16*, 9285−9298.

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

Wang, J.-F., & Cech, T. R. (1992) *Science 256*, 526−529.

Wang, J.-F., & Cech, T. R. (1994) *J. Am. Chem. Soc. 116*, 4178−4182.

Warshaw, M. M., & Cantor, C. R. (1970) *Biopolymers 9*, 1079−1103.

Westman, E., & Strömberg, R. (1994) *Nucleic Acids Res. 22*, 2430−2431.

Westman, E., Stawinski, J., & Strömberg, R. (1993) *Collect. Czech. Chem. Commun. 58*, 236.

Westman, E., Sigurdsson, S., Stawinski, J., & Strömberg, R. (1994) presented at the 21th Symposium on Nucleic Acids, Matsuyama, Japan, *Nucleic Acids Res., Symp. Ser. 31*, 25−26.

Whoriskey, S. K., Usman, N., & Szostak, J. W. (1995) *Proc. Natl. Acad. Sci. U.S.A. 92*, 2465−2469.

Zieba, K., Chu, T. M., Kupke, D. W., & Marky, L. A. (1991) *Biochemistry 30*, 8018−8026.

BI9523466