FTIR Spectroscopic Studies of Oligonucleotides That Model a Triple-Helical Domain in Self-Splicing Group I Introns†

Munna Sarkar,‡ Utz Dornberger,§ Eriks Rozners,|| Hartmut Fritzsche,§ Roger Strömberg||,⊥ and Astrid Gräslund*,‡

Department of Biophysics, Arrhenius Laboratories, Stockholm University, S-106 91 Stockholm, Sweden, Faculty of Biology and Pharmacy, Institute of Molecular Biology, Friedrich Schiller University, Jena, Winzerlaer Strasse 10, D-07745 Jena, Germany, Department of Organic Chemistry, Arrhenius Laboratories, Stockholm University, S-106 91 Stockholm, Sweden, and Laboratory of Organic and Bioorganic Chemistry, Department of Medical Biochemistry and Biophysics, Karolinska Institute, S-171 77 Stockholm, Sweden

Received January 30, 1997; Revised Manuscript Received September 24, 1997⊗

ABSTRACT: Fourier Transform infrared (FTIR) spectroscopy was used to characterize the Mg^{2+} dependent association of a 23-mer mixed ribo-deoxyribonucleotide (23-mer RNA) and a 7-mer oligoribonucleotide (7-mer RNA) that models the *triple-helical domain* of a self-splicing group I intron [Sarkar et al. (1996) *Biochemistry 35*, 4678−4688]. To elucidate the effect of deoxyribose substitution in the entire backbone, as well as at specific positions, in the assembly of the triple-helical domain, parallel studies were carried out on the association of pure deoxyribonucleotides having base sequences corresponding to the oligoribonucleotides and also between 23-mer RNA and two 7-mer RNA variants. In the variants, either the ribose attached to G451 or the ribose attached to U453 was changed to deoxyribose. FTIR-monitored thermal denaturation of the two 23-mer hairpins shows two distinct melting regions in 1 M NaCl, in case of the RNA hairpin but not for the 23-mer DNA. Triple-helix association between the two strands (7-mer and 23-mer) studied by FTIR show that only when both strands are RNA, association takes place with the formation of the P6 helix. Our results also show that the interactions between the two RNA strands involve some participation of the riboses, which could also involve the 2′-OH groups of the RNA backbone. The assembly of the triple-helical domain is not possible with a deoxyribose backbone and is completely perturbed even when only one ribose at either G451 or U453 position is substituted by deoxyribose.

RNA molecules with only four nucleotide building blocks are capable of adopting diverse three-dimensional structures that are essential to their function. The diversity of structures is achieved by making use of some preferred motifs and their combinations. Identifying these preferred motifs or building blocks and the interplay of different factors that determine the structure of these motifs is therefore of fundamental importance. The triple-helical domain in the catalytic conserved core of the self-splicing group I intron is one such motif. Nucleoside triple interactions involving major and minor grooves between paired regions P4 and P6 and single-stranded joining regions J6/7 and J3/4, respectively, have been proposed to constitute the triple-helical domain in the *Tetrahymena thermophila* intron (Michel et al., 1989, 1990; Michel & Westhof 1990). It occurs at the interface of two domains of coaxially stacked helices that are juxtaposed to create the active core (P4−P6 and P3−P9) and help in the correct orientation of these two domains (Doudna & Cech, 1995). A recently solved crystal structure of a 160-nucleotide P4−P6 domain of the *Tetrahymena* intron (Cate et al., 1996) showed that the overall structure of the triple-helical domain agrees with the Michel and Westhof model (1990) that predicted the J3/4 and J6/7 strands to lie in the minor and major grooves of P6 and P4 helices, respectively, even though the details of the structure of this domain were not well resolved. The NMR study of a corresponding 25-mer oligoribonucleotide that models the junction of P4−P6 helices and only the J3/4 single-stranded region also showed that the J3/4 single stranded region was associated as a triplet-forming strand in the minor groove of P6 (Chastain & Tinoco, 1992, 1993). However this is in contrast to what was observed in the NMR study of a 34 nucleotide RNA molecule that was designed to model the three-dimensional structure of both the J3/4 and J6/7 single-stranded regions and the P4−P6 junction of group I introns (Nowakowski & Tinoco, 1996). This shows that the modeling of a whole triple-helical domain by a single-stranded oligonucleotide can be ambiguous.

The present study concerns another approach using two oligonucleotides that could model the tertiary interactions involved in the triple-helical domain at the conserved active site of the self splicing group I intron from the bacteriophage T4 *nrdB* pre-mRNA. The T4 *nrdB* intron lacks the P5 bc subdomain of the *Tetrahymena* intron which makes several important tertiary contacts with P4−P6 that stabilize the compact architecture of the P4−P6 domain (Cate et al., 1996). A 23-mer RNA was synthesized as a mixed oligoribo-deoxyribonucleotides containing the J3/4-P4 (5′-end bases of P4) (nucleotides 42−51 in the original intron sequence)

†This study is supported by grants from the Swedish Natural Science Research Council, the European Commission Contract BMH4-CT96-0848, the Swedish Institute, the Magnus Bergwall Foundation, the Carl Trygger Foundation, the Deutscher Akademischer Austauschdienst (DAAD), and the Thuringian Ministry of Science, Research and Culture.

* Author to whom correspondence should be addressed.
‡ Department of Biophysics, Stockholm University.
§ Friedrich Schiller University.
|| Department of Organic Chemistry, Stockholm University.
⊥ Karolinska Institute.

⊗ Abstract published in *Advance ACS Abstracts,* November 15, 1997.

S0006-2960(97)00224-9 CCC: $14.00

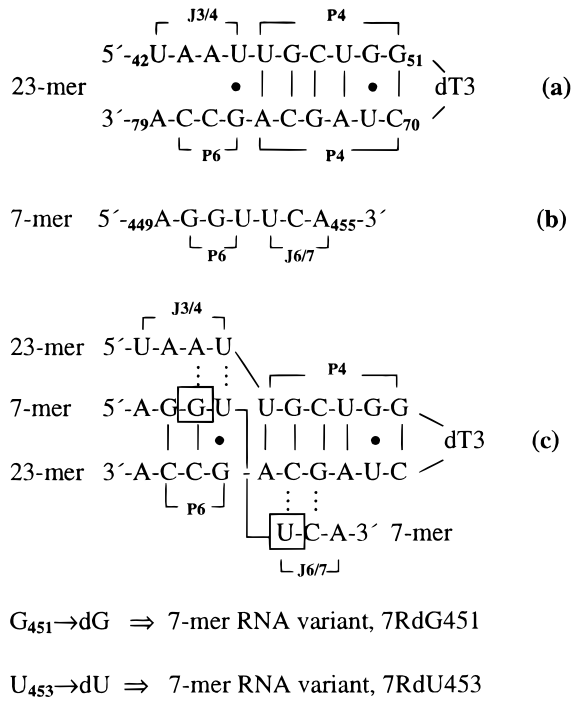

FIGURE 1: (a) Model sequences of 23-mer RNA (nucleotides 42−51 and 70−79; P5 loop exchanged for dT3, i.e., for nucleotides 52−69) (b) 7-mer RNA (nucleotides 449−455). (c) Schematic representation of the possible interactions involved in the association of 23-mer and 7-mer RNA. Boxes indicate the positions where the riboses have been changed to deoxyribose sugar in the two 7-mer RNA variants 7RdG451 and 7RdU453.

followed by three deoxythymidines (dT3) replacing the P5 loop and 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 is replaced by dT3 since they form a stable hairpin loop (Sarkar et al., 1996) (Figure 1a). A heptaribonucleotide (7-mer RNA) containing the 3′-end bases of P6 and J6/7 (nucleotides 449−455), with which a triplex has been shown to form (Sarkar et al., 1996), was synthesized as a pure oligoribonucleotide (Figure 1b). The nucleotide numbers used in the text correspond to those in the original sequence. In our previous study using UV and circular dichroism spectroscopy (Sarkar et al., 1996) we have shown that 23-mer RNA forms a stable hairpin modeling the P4 base-paired region (Figure 1a) and that the 23-mer RNA hairpin and 7-mer RNA single strand associate in the presence of Mg^{2+} to form a "plaited triple helix" modeling the triple-helical domain of the intron core. Figure 1c is a schematic representation of some possible interactions between the 23-mer and 7-mer RNA. One advantage of this two-strand model system over a single-strand model system is that the structure formed is not constrained by the covalent connectivity in a short strand. Hence the assembly process in the present case may in certain aspects better resemble the native folding.

Fourier Transform infrared spectroscopy (FTIR) has proven to be well suited for the characterization of nucleic acid conformation (Taillandier & Liquier, 1992) providing information on interactions involving specific groups. Whereas several studies have been devoted to the application of IR spectroscopy to polyribonucleotides (Tsuboi, 1969) and native RNA (Li et al., 1984) there exist only a few reports in which it has been applied to characterize RNA oligomers (Liquier et al., 1991, 1995; Klinck et al., 1994). In this study we have used FTIR spectroscopy to characterize the association between the 23-mer and 7-mer RNAs that model the triple-helical domain in the T4 *nrdB* pre-mRNA intron.

In order to highlight the effect of backbone and/or sugar conformation in general, deoxyribose nucleotides having the base sequences corresponding to the 23-mer RNA (23-mer DNA) and 7-mer RNA (7-mer DNA) have also been studied. In the DNA strands, uracil has been replaced by thymidine. FTIR monitored thermal melting of the two 23-mer hairpins (RNA and DNA) show interesting structural differences between the two. The association between the 23-mer RNA and two 7-mer RNA variants was also characterized. In one variant, the G451 in the original sequence has a deoxyribose sugar and in the other U453 has a deoxyribose sugar. The G451 is proposed to be involved in a Watson−Crick base pair (C−G) in the P6 helix region of our model along with a P6-J3/4 nucleoside triple interaction in the minor groove, whereas U453 is proposed to be involved in a nucleoside triple interaction in the major groove (Michel & Westhof, 1990). The positions of the two deoxyribose substitutions are equidistant from the 5′-end and 3′- end of the 7-mer RNA, respectively (Figure 1c). The studies with the variants of 7-mer RNA have been carried out to elucidate the importance of sugar conformation and/or 2′-OH groups of the ribose sugar in two strategic positions for the formation of the triple-helical domain.

MATERIALS AND METHODS

Oligonucleotides

The mixed ribo-deoxyribonucleotide (23-mer RNA), the oligoribonucleotide (7-mer RNA), and the 7-mer RNA variants, 7RdG451 and 7RdU453, were synthesized and purified as described previously (Sarkar et al., 1996). The monomeric building blocks for the incorporation of deoxyuridine and deoxyguanosine were prepared as follows.

5′-O-(4-Monomethoxy)trityl-2′-deoxyuridine-3′-hydrogenphosphonate triethylammonium salt was prepared from 2′-deoxyuridine (Janssen Chimica) using the procedures reported for 5′-*O*-(4,4′-dimethoxy)tritylation (Connolly, 1991) and the synthesis of 5′-*O*-(4,4′-dimethoxy)trityl-2′-*O*-*tert*-butyldimethylsilyluridine 3′-*H*-phosphonate [as in Stawinski et al. (1988) except that dichloromethane was used as solvent and the coevaporation with triethylamine/pyridine (1:3) was omitted], R_f = 0.15 ($CHCl_3$/methanol, 4:1, v/v) and 0.33 (isopropanol/water/25% aq. NH_3, 85:5:10, v/v/v). ^{1}H NMR: 8.27 (s, 1H, NH), 7.72 (d, J = 8.2 Hz, 1H, H6), 7.39−7.11 and 6.82 (14H, Ar), 6.96 (d, J = 625 Hz, 1H, PH), 6.30 (m, 1H, H1′), 5.29 (d, 1H, H5), 5.02 (m, 1H, H3′), 4.24 (m, 1H, H4′), 3.76 (s, 3H, OCH_3), 3.43 (m, 2H, H5′), 3.04 (q, J = 7.3 Hz, 6H, CH_2N), 2.61 and 2.35 (2m, 2H, H2′), 1.32 (t, 9H, CH_3).

5′-O-(4-Monomethoxy)trityl-N-phenoxyacetyldeoxyguanosine-3′-hydrogenphosphonate triethylammonium salt was prepared from *N*-phenoxyacetyldeoxyguanosine (Schulhof et al., 1987) using reported procedures for 5′-*O*-(4,4′-dimethoxy)tritylation (Connolly, 1991) and the synthesis of 5′-*O*-(4,4′-dimethoxy)trityl-2′-*O*-*tert*-butyldimethylsilyluridine 3′-H-phosphonate [as in Stawinski et al. (1988) except that dichloromethane was used as solvent and the coevaporation with triethylamine/pyridine (1:3) was omitted], R_f = 0.22 ($CHCl_3$/methanol, 4:1, v/v) and 0.33 (isopropanol/water/25%

aq NH_3, 85:5:10, v/v/v). ^{1}H NMR: 8.07 (s, 1H, NH), 7.86 (s, 1H, H8) 7.38−6.77 (19H, Ar), 6.90 (d, J = 625 Hz, 1H, PH), 6.28 (m, 1H, H1′), 5.11 (m, 1H, H3′), 4.64 (s, 2H, $COCH_2O$), 4.33 (m, 1H, H4′), 3.74 (s, 3H, OCH_3), 3.35 (m, 2H, H5′), 3.05 (q, J = 7.3 Hz, 6H, CH_2N), 2.86−2.70 (m, 2H, H2′), 1.29 (t, 9H, CH_3).

Oligodeoxyribonucleotides were obtained from Kebo Lab, Stockholm, Sweden (synthesized according to the phosphotriester method and purified by HPLC). They were further purified by passing the samples through NAP (Pharmacia) columns and freeze-dried from water.

The oligonucleotide concentration in each case, was determined spectroscopically for a solution that was diluted 10^3 times the actual concentration used for the FTIR experiments. Single-strand absorbance at 25 °C was determined by the extrapolation of the upper base line in the UV melting curve measured at 260 nm. The corresponding 25 °C extinction coefficients in single strands were calculated by nearest neighbor analysis (Puglisi & Tinoco, 1989). The following extinction coefficients in single strands at 260 nm and 25 °C were used: 23-mer RNA (2.28×10^5 M^{-1} cm^{-1}); 7-mer RNA, 7RdG451 and 7RdU453 (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}). Stock solution concentrations varied from 50 to 100 mM in bases. The final sample solutions contained either 10 mM sodium cacodylate buffer, 0.1 mM EDTA, or 20 mM sodium cacodylate buffer, 0.2 mM EDTA, with pH in the range 6.8−7. NaCl or $MgCl_2$ was added as required. The pH values given in the text for D_2O solutions are the pH values measured in the corresponding H_2O solutions. For preparation of deuterated samples, the sample solutions were evaporated to dryness under a gentle flow of dry nitrogen and then redissolved in the same volume of D_2O.

Fourier Transform Infrared Spectroscopy

FTIR measurements were performed using a Bruker IFS-66 spectrometer equipped with a DGTS detector. For melting or thermal denaturation experiments about 7 μL sample solution of either 23-mer RNA or DNA in 10 mM sodium cacodylate buffer, 0.1 mM EDTA in D_2O (pH = 7.0) was placed in a demountable temperature-controlled liquid cell (Harrick) with CaF_2 windows. The temperature was measured with a sensor attached directly to the windows. The path length was 56 μm for D_2O buffer. The sample temperature was increased at a rate of 0.7 °C/min while recording 32 interferograms per minute. The temperature bath was controlled directly by the FTIR spectrometer software from Bruker.

For the association reaction between 23-mer hairpins and 7-mer single strands the mixture containing 1:1 stoichiometry in single strands was first heated to 50 °C and then slowly annealed. This was done to facilitate complex formation. About 1−1.5 μL droplets of concentrated (8−13 mM in single strands) of RNA or DNA samples were deposited in water-jacketed cells sealed by ZnSe windows. The sample temperature was kept ∼10 °C by circulating cooled water through the jacketed cells. The sample chamber of the spectrometer was continuously purged with dry air. FTIR spectra were recorded using identical scanning parameters and temperatures. The resolution was set to 2 cm^{-1}, and 32 interferograms were accumulated, coadded, and Fourier transformed using the Happ Genzel apodization function. Data treatment was limited to solvent background subtraction and baseline correction. Care was taken to maintain identical experimental conditions and concentrations of the samples in recording the spectra of the 23-mer RNA, the 7-mer RNA, and the 1:1 stoichiometric mixture of the two. This enabled us to directly compare the added spectrum of the component strands with that of the observed spectrum of the mixture. The phosphate symmetric stretching vibration around 1083 cm^{-1} is often used as an internal standard for spectral normalization (Akhebat et al., 1992). In all cases, where the phosphate symmetric stretching vibrations remained unchanged upon mixing the 23-mer and 7-mer strands, a complete overlap of this band was observed between the added spectra of the components and that of the mixture. This showed that no additional spectral normalization was required.

RESULTS

Thermal Denaturation of 23-mer RNA and DNA Hairpins: In-Plane Double-Bond Base Vibrations (1800−1500 cm^{-1})

The FTIR spectra of nucleic acids in the region of 1800−1500 cm^{-1} consist of absorption bands originating from the in-plane double-bond vibrations of the bases. These bands are extremely sensitive to base interactions involving hydrogen bonding, such as base pairing. This spectral region was studied in D_2O due to the presence of the strong interfering water absorption band around 1650 cm^{-1}. Figure 2 presents the FTIR spectra recorded at low and high temperatures of RNA hairpin (Figure 2a) and DNA hairpin (Figure 2b) in the presence of 1 M NaCl. Two intense bands around 1686 and 1652 cm^{-1} are observed in the spectrum of 23-mer RNA hairpin at 15 °C (Figure 2a). The band at 1686 cm^{-1} has been assigned to the C6═O6 stretch of base-paired guanines plus C2═O2 bond stretching vibration of uracils and thymines. The 1652 cm^{-1} band has been assigned to C2═O2 of cytosines plus C4═O4 stretching vibrations of uracils and thymines (Tsuboi, 1969; Taillandier & Liquier, 1992). Upon melting, the band around 1686 cm^{-1} decreases in intensity and an intense band at 1658 cm^{-1} appears (Figure 2a). This band is due to the overlapping contribution of the stretching vibrations of C6═O6 of free or non-base-paired guanines, the C2═O2 of free cytosines, and the C4═O4 of free thymines or uracils (Tsuboi, 1969; Taillandier & Liquier, 1992; Dagneaux et al., 1994). The strong band around 1624 cm^{-1} which increases in intensity and shifts to 1621 cm^{-1} on melting, can mainly be assigned to the C═C and C═N stretching vibrations of adenine along with a minor contribution of some ring vibrations of free cytosines (Tsuboi, 1969; Dagneaux et al., 1994). The band at 1575 cm^{-1} which increases in intensity on melting of the hairpin (Figure 2a) arises mainly from the guanine C═N stretching vibrations.

The FTIR spectrum of the 23-mer DNA hairpin at 15 °C (Figure 2b) shows three bands between 1700 and 1630 cm^{-1}. The band at 1692 cm^{-1} is assigned to the carbonyl stretching vibrations of guanines in the duplex part of the hairpin and a minor contribution of C2═O2 of thymines. The broad band centered around 1665 cm^{-1} involves contributions of C2═O2 of cytosines and the C4═O4 stretching vibrations of thymines. The band at 1644 cm^{-1} is due to C═C, C═N

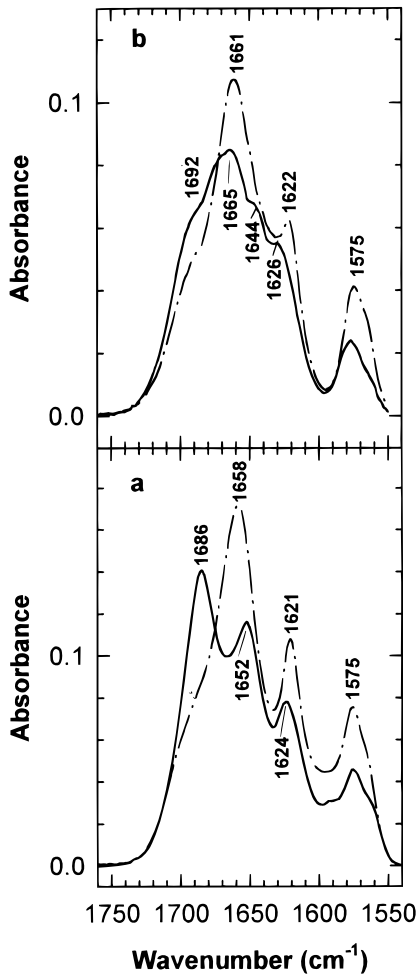

FIGURE 2: (a) FTIR spectra of 23-mer RNA hairpin in the spectral region of 1760−1540 cm^{-1} in 10 mM sodium cacodylate buffer, 0.1 mM EDTA, pH = 6.9, and 1 M NaCl in D_2O at 15 °C (—) and 86 °C (− · −). Strand concentration = 1.54 mM. (b) FTIR spectra of 23-mer DNA hairpin in the same spectral region in 10 mM sodium cacodylate buffer, 0.1 mM EDTA, pH = 6.9, and 1 M NaCl in D_2O at 15 °C (—) and 86 °C (− · −). Strand concentration = 1.42 mM.

ring vibrations of thymines engaged in Watson−Crick base pairing (Tsuboi, 1969; Dagneaux et al., 1995).

The thermal melting of the 23-mer RNA and DNA hairpins were followed by the changes in the (i) band intensity ratio 1658/1686 cm^{-1} for 23-mer RNA and 1661/1692 cm^{-1} for 23-mer DNA, (ii) intensity of the 1621−1625 cm^{-1} adenine band, and (iii) intensity of the 1575 cm^{-1} guanine band.

Whereas (i) should reflect the melting of all the base pairs in general, (ii) and (iii) should reflect the changes in the adenine- and guanine-rich regions respectively.

Figure 3 shows the melting curves of 23-mer RNA and DNA hairpins in the presence of 1 M NaCl in D_2O buffer and the corresponding first derivatives of the curves, monitored by the changes in different IR band intensities. Figure 3a displays the melting of 23-mer RNA monitored by the changes in the intensity ratio 1658/1686 cm^{-1}. The melting curve exhibits a biphasic behavior with transitions at 38.9 and 71.5 °C. The higher transition temperature (t_m = 71.5 °C) is similar to that observed in UV melting curve of the same 23-mer RNA hairpin in 1 M NaCl at a much lower strand concentration of 5.2 μM. Hence this transition is attributed to the monomolecular melting of the 23-mer RNA hairpin. Figure 3b shows the melting of the 23-mer DNA monitored by the changes in the band intensity ratio 1661/1692 cm^{-1}. The 23-mer DNA displays one major transition at 66.3 °C. This transition temperature is similar to that observed in the UV melting of the same 23-mer DNA in 1 M NaCl but at a much lower strand concentration (4 μM) (data not shown), which indicates that this transition is due to the monomolecular melting of 23-mer DNA hairpin. Due the noise level in this IR melting curve the apparent transition around 30 °C is not significant. It should be mentioned that in the absence of added salt, both 23-mer RNA and 23-mer DNA hairpins show monophasic melting transitions when monitored by the band intensity ratios 1658/1686 cm^{-1} and 1661/1690 cm^{-1}, respectively (data not shown).

To find out whether the lower RNA transition (t_m = 38.9 °C) (Figure 3a) represents melting of intermolecular duplex that may form by the association of two 23-mer RNA molecules, the melting was also monitored by the changes in the adenine and the guanine bands. Figure 3c shows the melting curve monitored at the 1621 cm^{-1} adenine band. The melting curve exhibits a biphasic transition with the lower-temperature transition much more pronounced relative to the high-temperature transition than was observed in Figure 3a. This indicates that the 38.9 °C transition probably reflects the melting of an adenine-rich region. Figure 3d shows the same transition monitored at the 1575 cm^{-1} guanine band. Here the low-temperature transition (t_m = 36 °C) is very weak indicating that the lower-temperature transition does not reflect the melting of guanine-rich regions. From the base sequence of 23-mer RNA (Figure 1a), it is evident that the most probable intermolecular duplex that may form by the association of two 23-mer RNA molecules should essentially have ten base pairs involving guanines and four base pairs involving adenines with an internal loop of six thymines in the middle. The two U−A at the 5′- and 3′-ends are not considered to form base pairs. The melting of the intermolecular duplex should be characterized by base pair opening/destacking of more guanines than adenines and hence is expected to be more pronounced when monitored by a guanine band than by an adenine band. Hence for the low-temperature transition in Figure 3a to represent the melting of intermolecular duplex, it is expected to be more pronounced when monitored by the guanine band at 1575 cm^{-1} than when monitored by the adenine band at 1621 cm^{-1}. This is in contrast to our obtained results (Figure 3c,d), indicating that the low-temperature transition does not represent the disruption of an intermolecular helix between two 23-mer RNA molecules. In addition, a decrease in strand concentration from 1.54 mM to 0.9 and 0.4 mM did not result in significant shift in the low temperature transition with calculated t_m values within the experimental error range of ±3 °C. The melting was also monitored by the UV absorption band at 260 nm. In contrast to the IR melting curves, the low-temperature transition showed up as a very weak transition around 35 °C (limited by the experimental noise level) both at single strand concentration of 0.4 mM and also at a much lower concentration of 5.2 μM. The difference in the IR and UV melting profiles indicates that this transition reflects a stronger change in the hydrogen bonding network (as probed by IR) than a change in the stacking interaction between the bases (as probed by the UV absorption at 260 nm). Since the complete melting of

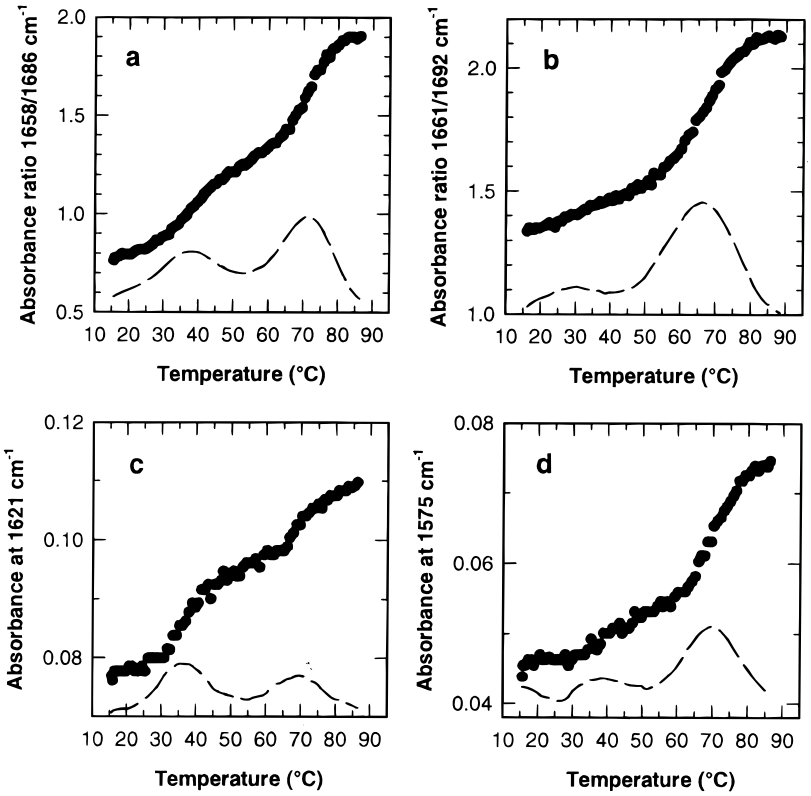

FIGURE 3: Melting curves (●) and the corresponding first-derivative spectra (− −) of 23-mer RNA hairpin (1.54 mM in strands) monitored by the changes in band intensity ratio 1658/1686 (a) and the 23-mer DNA hairpin (1.42 mM in strands) by the changes in the band intensity ratio 1661/1692 cm^{-1} (b); for the 23-mer RNA adenine band at 1621 cm^{-1} (c); and for the guanine band at 1575 cm^{-1} (d).

intermolecular duplex would involve breaking of several base pairs and destacking, it is expected to show up as a prominent transition also in the UV melting. All these results together indicate that the low-temperature transition is not due to melting of intermolecular duplex formed between two 23-mer RNA molecules; rather, it reflects the disordering of an adenine-rich region in the stem duplex of the 23-mer RNA hairpin.

The higher-temperature transition in Figure 3a mainly represents the melting of a guanine-rich region of the hairpin stem. In the case of the 23-mer DNA, the melting transition shows identical behavior as in Figure 3b irrespective of whether the changes in the adenine or guanine bands were monitored, which shows that in case of 23-mer DNA hairpin, the adenine- and guanine-rich regions melt together. Taken together our results indicate that only in case of the 23-mer RNA hairpin in high salt does the disordering of the adenine rich region of the stem duplex (Figure 1a) show up as a separate transition.

Association between 23-mer RNA/DNA Hairpins and 7-mer RNA/DNA Single Strands

The associations between the 23-mer hairpins and 7-mer single strands of both RNA and DNA were followed by FTIR spectroscopy in D_2O and H_2O buffer in the presence of 200 mM $MgCl_2$. The FTIR spectra of a 1:1 (in single strands) mixture of 23-mer hairpins and 7-mer single strands were compared with the "added" spectra obtained by adding the FTIR spectra of equimolar concentrations of 23-mer RNA/DNA hairpins and 7-mer RNA/DNA single strands, recorded separately under the same conditions of 20 mM sodium cacodylate buffer and 200 mM $MgCl_2$ at 10 °C, pH = 6.9. Care was taken to maintain identical experimental conditions and concentration of the samples.

Base Frequency Region: In-Plane Double-Bond Vibrations (1800−1500 cm^{-1})

The in-plane double-bond base vibrations of nucleic acids ocurring between 1800 and 1500 cm^{-1} are sensitive to the changes in base pairing and/or stacking which might occur when the two partner strands associate (23-mer and 7-mer RNA/DNA). Figure 4a shows the FTIR spectra of 7-mer RNA (i), 23-mer RNA (ii), the added spectra of equimolar concentrations of 23-mer and 7-mer RNA (iii), and 1:1 mixture in 200 mM $MgCl_2$ (iv).

The FTIR spectrum of 7-mer RNA showed a broad band centered around 1705 cm^{-1} (Figure 4a, i). This band may have contributions either from C2═O2 stretching vibrations of the uracils or from autoassociated guanines (Dagneaux et al., 1994) that could exist in intermolecular complexes which might form under the studied conditions. Second derivative of the FTIR spectrum showed this broad band to be composed of two overlapping bands one at 1705 and the other at 1692 cm^{-1}. To distinguish between the contributions from the uracils and the autoassociated guanines the experiment was repeated under identical conditions but with a fourteen fold decrease in sample concentration (14 to 1 mM). The spectral profile remained unaffected by the decrease in the strand concentration, strongly indicating the absence of any intermolecular complexes (data not shown). Even in the absence of magnesium, which is known to induce higher order structures in RNA, a 14-fold decrease in strand concentration did not affect the spectral profile of 7-mer RNA. To verify that the bands around 1705 cm^{-1} may in fact be due to uracil carbonyl vibrations and not to autoas-

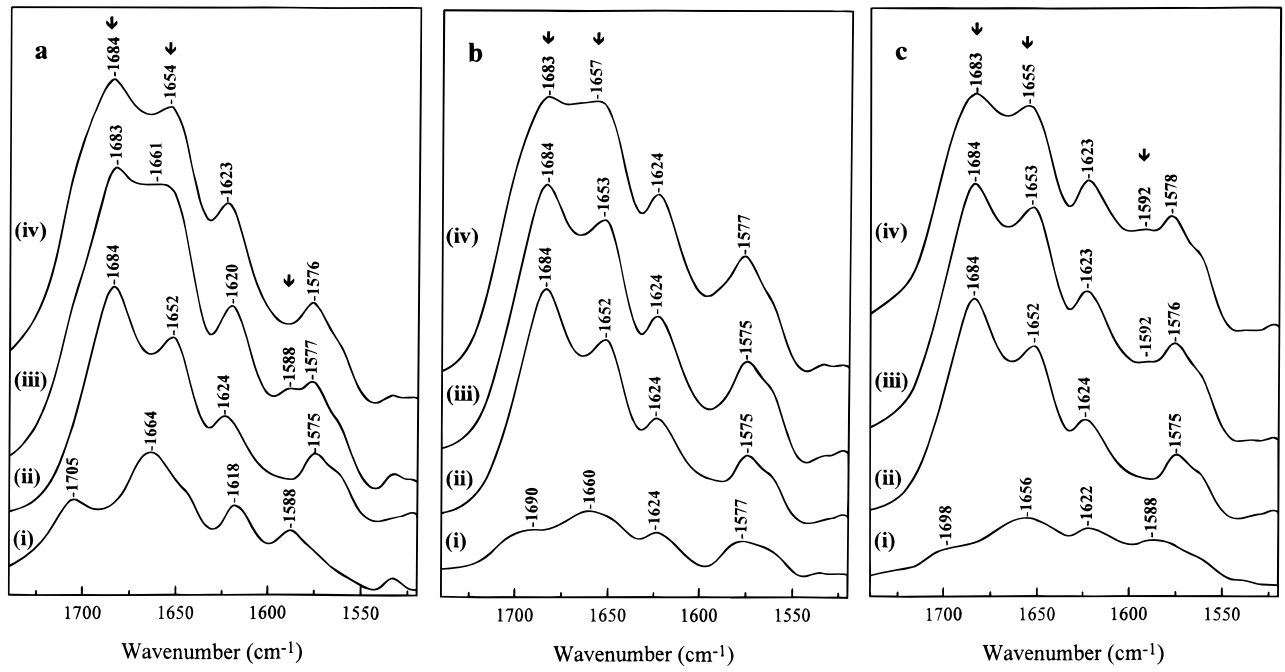

FIGURE 4: (a) FTIR spectra of the 1740−1520 cm^{-1} region of (i) 7-mer RNA, (ii), 23- mer RNA, (iii) "added" spectrum of 23-mer and 7-mer RNA (see text), and (iv) observed spectrum of 1:1 complex of 23-mer and 7-mer RNA in 20 mM sodium cacodylate, 0.2 mM EDTA, and 200 mM $MgCl_2$ in D_2O at 10 °C. Individual strand concentration is 10 mM. (b) FTIR spectra of the same spectral region in the same D_2O buffer of (i) 7RdG451 variant, (ii), 23-mer RNA, (iii) "added" spectrum of 23-mer RNA and 7RdG451 (see text), and (iv) the observed spectrum of 1:1 complex of the same. Individual strand concentration is 10 mM. (c) FTIR spectra of (i) 7RdU453 variant, (ii), 23-mer RNA, (iii) "added" spectrum of 23-mer RNA and 7RdU453 (see text), and (iv) the observed spectrum of 1:1 complex of the same. Individual strand concentration is 8 mM.

sociated guanines we recorded the FTIR spectrum of single stranded polyU in 10 mM sodium cacodylate buffer 0.1 mM EDTA, pH = 7.0, with 50 mM $MgCl_2$ in D_2O. A broad band was observed centered around 1700 cm^{-1} and was also found to be composed of two overlapping bands at 1705 and at 1692 cm^{-1} (as determined from the second derivative spectrum; data not shown) similar to that seen in the FTIR spectrum of 7-mer RNA recorded under identical conditions. We therefore assign this broad band around 1705 cm^{-1} in the 7-mer RNA spectrum to C2═O2 stretching vibrations of uracils that could be ordered or stacked in the single stranded form of the mixed base sequence of 7-mer RNA. Increase in temperature to 75 °C resulted in intensity decrease of the 1705 cm^{-1} band both in the spectra of 7-mer RNA and polyU reflecting a disordering in the single stranded form. Furthermore the presence of the strong broad band around 1664 cm^{-1} assigned to an overlapping contribution of C6═O6 of free or non-base-paired guanines, C4═O4 of free uracils and C2═O2 of free cytosine stretching vibrations along with the 1618 cm^{-1} band mainly assigned to C═C, C═N ring vibrations of free cytosine (Tsuboi, 1969; Dagneaux et al., 1994) are in complete agreement with the 7-mer RNA being essentially in single-stranded form. The band at 1588 cm^{-1} is mainly due to the strong C═N ring vibrations of guanine (Dagneaux et al., 1994). Taken together these observations confirm that the 7-mer RNA had not autoassociated under the experimental conditions employed in this study.

The observed spectrum of the 1:1 mixture of 23-mer RNA and 7-mer RNA (Figure 4a, iv) is significantly different from that of the added spectrum (Figure 4a, iii). In all cases, the 1:1 mixture of 23-mer RNA hairpin and 7-mer RNA single strand was heated to 50 °C and then slowly annealed to facilitate complex formation. Without preheating, no complex was formed and no significant difference between the added and the complex spectra was observed. The intensity ratio of the 1684/1654 cm^{-1} bands which is larger for a larger number of base pairs (Tsuboi, 1969) is higher in the observed spectrum (Figure 4a, iv) compared to the added spectrum (Figure 4a, iii). This indicates that association of 23-mer RNA and 7-mer RNA occurs in the 1:1 mixture, with the formation of additional base pairs. Another important observation is the absence of the 7-mer RNA guanine band at 1588 cm^{-1} in the complex spectrum. This suggests that the two guanines of the 7-mer RNA are changed in stacking geometry in the complex which could be a consequence of Watson−Crick base pairing between the two guanines and two cytosines of the 23-mer RNA (Figure 1c). In the corresponding experiments with DNA oligomers (data not shown) no change in the 1578 cm^{-1} guanine band was observed in the complex spectrum relative to the added spectrum.

Figure 4b shows the FTIR spectra of the 7-mer RNA variant 7RdG451 (i), 23-mer RNA (ii), added spectrum of equimolar concentrations of 23-mer RNA and 7RdG451 (iii), and the 1:1 mixture of the same in 200 mM $MgCl_2$ (iv). The intensity ratio of the 1683/1657 cm^{-1} bands in the 1:1 mixture (Figure 4b, iv) is lower than the added spectrum (Figure 4b, iii), which is in contrast to that observed in the association between 23-mer and 7-mer RNA (Figure 4a, iii and iv). This could imply loss of H-bonding, and/or base pairing in the 23-mer RNA itself on heating and subsequent annealing of the 23-mer RNA hairpin, in the presence of 7RdG451 in 1:1 mixture of the same. Similar differences in the band intensity ratio 1683/1655 cm^{-1} was seen between the spectra of 1:1 mixture of 23-mer RNA and 7-mer RNA variant 7RdU453 in 200 mM $MgCl_2$ and the corresponding added spectrum of the components (Figure 4c). The exist-

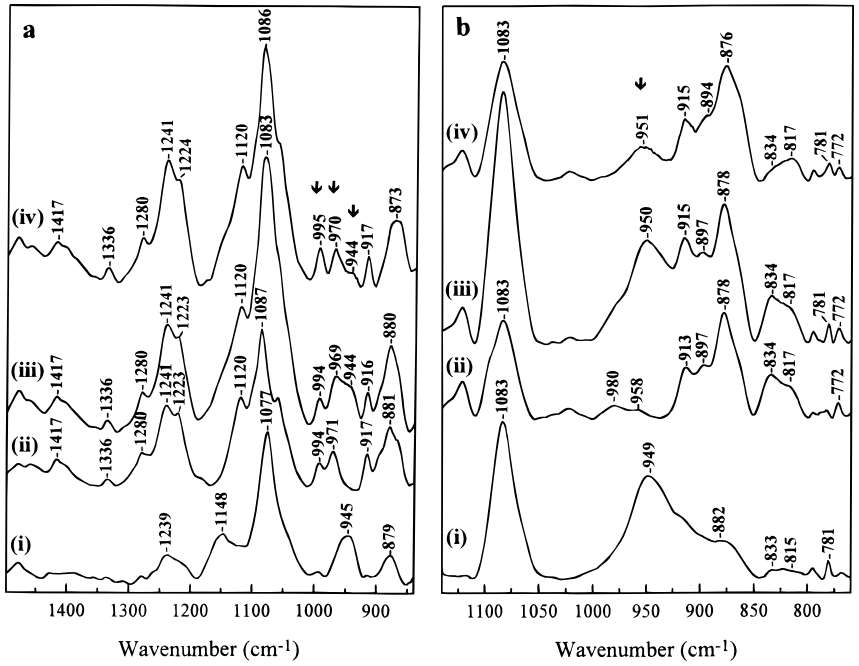

FIGURE 5: (a) FTIR spectra of the 1500−840 cm^{-1} region of (i) 7-mer RNA, (ii) 23-mer RNA, (iii) "added" spectrum of 23-mer and 7-mer RNA (see text), and (iv) the observed spectrum of 1:1 complex of 23-mer and 7-mer RNA in 20 mM sodium cacodylate, 0.2 mM EDTA, and 200 mM $MgCl_2$ in H_2O at 10 °C. Individual strand concentration is 10 mM. (b**)** FTIR spectra of the 1140-760 cm^{-1} region of (i) 7-mer RNA, (ii) 23-mer RNA, (iii) "added" spectrum of 23-mer and 7-mer RNA (see text), (iv) observed spectrum of 1:1 complex of 23-mer and 7-mer RNA in 20 mM sodium cacodylate, 0.2 mM EDTA, and 200 mM $MgCl_2$ in D_2O at 10 °C. Individual strand concentration is 10 mM.

ence of the 7-mer guanine band at 1588 cm^{-1} both in the added spectrum of the components (Figure 4c, iii) and in the mixture (Figure 4c, iv) (1592 cm^{-1}) suggests that the two guanines of the variant 7RdU453 are not involved in base pairing. This is in contrast to that observed in the association between the 23-mer RNA and 7-mer RNA (Figure 4a).

Backbone−Sugar Vibrational Frequency Region (1500−750 cm^{-1}) in H_2O and D_2O

Figure 5 shows the FTIR spectra of RNA oligomers (23-mer and 7-mer) and the corresponding added and observed spectra of 1:1 mixture in H_2O (Figure 5a) and in D_2O (Figure 5b) solutions in the backbone−sugar frequency region. The bands of sugar, base−sugar, and backbone vibrations are between 1500 and 1200 cm^{-1}. They appear at the expected wavenumbers for the C3′-endo anti conformation of riboses characteristic of A-form geometry (Taillandier & Liquier, 1992; Quali et al., 1993; Taboury et al., 1985) and do not change significantly upon complex formation (Figure 5a, iii and iv). The band around 1083 cm^{-1} in the RNA spectrum in H_2O (Figure 5a) has been assigned to the symmetric PO_2^- stretching mode coupled with the C5′−O5′vibrations (Taillandier et al., 1985). This band is unexpectedly intense in the single stranded 7-mer RNA spectrum (1077 cm^{-1}) compared to the antisymmetric stretching vibrational band around 1239 cm^{-1} (Figure 5a, i). In the complex spectrum (Figure 5a, iv), this band is narrowed relative to the added spectrum (Figure 5a, iii) and it is shifted to a higher wavenumber. This could indicate that all phosphates in the complex are more structurally equivalent than in the two components.

Several interesting changes in the ribose vibrations were observed in the spectral region between 1000 and 940 cm^{-1} both in the H_2O (Figure 5a) and in the D_2O (Figure 5b) buffer. The two bands at 995 and 970 cm^{-1} in the complex spectrum (Figure 5a, iv) occur at similar positions but with different relative intensities compared to those observed in the 23-mer RNA hairpin alone (Figure 5a, ii). The band at 995 cm^{-1}, which is absent in D_2O solutions, has been tentatively assigned to a ribose vibrational mode involving the 2′-OH group (Liquier et al., 1991). The 970 cm^{-1} band has been assigned to the C4′-C5′H and O5′-C5′H vibrations (Letellier et al., 1989; Dohy et al., 1989).

In the 7-mer RNA spectrum, a strong band is observed at 945 cm^{-1} in H_2O (Figure 5a) and at 949 cm^{-1} in D_2O (Figure 5b) solutions and is completely absent in the 23-mer RNA hairpin. Upon complex formation this band is strongly reduced in intensity both in the H_2O (Figure 5a, iv) and D_2O (Figure 5b, iv) solutions. This intensity of the 945 cm^{-1} band is correlated to the intensity of the symmetric stretching PO_2^- band around 1083 cm^{-1}, and this correlation remains when the Mg^{2+} concentration or temperature is varied. Its exact position is dependent on both the Mg^{2+} concentration and the temperature. We tentatively assign the 945 cm^{-1} band to a ribose vibration coupled to the phosphate backbone. The change in the relative intensities of the 995 and 970 cm^{-1} bands and the decrease in intensity of the 945 cm^{-1} band on association of 23-mer and 7-mer RNA, reflect some interactions at the ribose level in the RNA complex that could also involve the 2′-OH groups.

The spectral region below 900 cm^{-1}, observed by us only in D_2O solutions (Figure 5b), contains well-known marker bands for the sugar conformation. The prominent bands at 876 and 817 cm^{-1} in the complex (Figure 5b, iv), indicate predominant C3′-endo conformation of the riboses. The band around 834 cm^{-1} in the 23-mer RNA (Figure 5b, ii), which is indicative of C2′-endo conformation (Taillandier et al., 1985), has a lower intensity in the complex than in the added spectrum (Figure 5b, iii and iv). The presence of

this band in the 23-mer RNA hairpin could indicate that the deoxythymidines in the loop are in the expected C2′-endo anti conformation.

No significant differences were detected between the added and the observed spectra of the 1:1 mixture of 23-mer RNA and the 7-mer variants, 7RdG451 and 7RdU453, in the 1500−750 cm^{-1} region. This was also true for the DNA oligomers.

DISCUSSION

We have shown by FTIR spectroscopy that both the 23-mer RNA and the 23-mer DNA form hairpins even at millimolar strand concentration, thus correctly modeling the P4 base-paired region. From Figure 1a it is evident that the stem duplex of the 23-mer hairpins consists of two regions, viz., a non-paired end of the stem rich in adenines and a base-paired region rich in guanines. FTIR-monitored thermal melting of 23-mer RNA and DNA hairpins brings out interesting differences in the melting profiles of these two hairpins at high salt concentration (1 M NaCl), which could not be resolved by UV melting. The 23-mer RNA hairpin shows a biphasic melting behavior with transitions at 38.9 and 71.5 °C in contrast to that observed for the 23-mer DNA hairpin and also to that observed for the 23-mer RNA hairpin under conditions of low salt concentration. The higher temperature transition represents mainly the melting of the guanine-rich base-paired region, whereas the low-temperature transition represents the disordering of the adenine rich non-paired end of the stem of the RNA hairpin, which is apparently ordered by hydrogen bonding and/or stacking in high salt concentration, that results its melting to show up as a separate transition. Under low salt concentrations, lack of additional stacking and/or hydrogen bonding in the adenine-rich non-paired end of the hairpin stem, results in the absence of the low-temperature transition in the FTIR-monitored melting profiles of RNA/DNA haiprins. The absence of the low-temperature transition in case of 23-mer DNA hairpin even in 1 M NaCl could indicate that the deoxyribose backbone is incapable of permitting the additional stacking and/or base pairing like the RNA hairpin in high salt. Possible reasons could be different stacking properties of A and B form helices and/or the lack of a hydrogen bond donor in the backbone like the 2′-OH groups of the RNA, which are known to make favorable contacts that stabilize RNA structure (Holland & Hoffman, 1996; Chastain & Tinoco, 1993; Cate et al., 1996).

FTIR spectroscopy allows us better insight than UV or CD spectroscopy (Sarkar et al., 1996) into the interactions involved in the association of the 23-mer RNA hairpin and the 7-mer RNA single strand. The proposed structure of the present model system of the triple-helical domain in group I intron requires base pairing between the 5′-end bases of the 7-mer RNA and 3′-end bases of the 23-mer RNA to model the P6 helix, as well as the nucleoside triple interactions between J3/4 and J6/7 with P6 and P4 helices, respectively (Figure 1c). The P6 helix involves the formation of the two C-G base pairs, where the 7-mer RNA guanines participate. The absence of the 7-mer RNA guanine band at 1588 cm^{-1} in the 1:1 complex spectrum compared to the added spectrum of 23-mer and 7-mer RNA, suggests that the two guanines of the 7-mer RNA are involved in base pairing in the complex. This is consistent with the model of the P6 helix. The absence of complexation between the DNA strands that involves formation of additional C−G base pairs points to the fact that the deoxyribose backbone is incapable of the association that correctly models the triple-helical domain in group I intron.

Interestingly, the complexation was completely perturbed when only one ribose was changed to deoxyribose at G451 or at U453 in the 7-mer RNA, as seen in the association between 23-mer RNA and the variants 7RdG451 and 7RdU453. In the proposed model, the G451 is involved, both in Watson−Crick base pairing to form the P6 helix and nucleoside triple interactions with J3/4, whereas U453 is involved in a nucleoside triple interaction. The absence of complexation in these cases could indicate either direct or indirect participation of 2′-OH groups at these positions (Chastain & Tinoco, 1992, 1993; Whoriskey et al., 1995). Indirect participation would mean that the absence of 2′-OH group could result in a different sugar conformation. This might affect helical parameters which in turn could hinder triplex formation by interfering with the proper positioning of base residues for hydrogen bond formation.

The association between the 23-mer and 7-mer RNA results in several changes in the ribose vibrations in the spectral region 1000−800 cm^{-1}. The change in the relative intensities of the two ribose bands at 995 and 970 cm^{-1} as well as the decrease in the intensity of the 945 cm^{-1} band indicate interactions involving the riboses in the complex that might also involve the 2′-OH groups in the RNA strands. The symmetric PO_2^- stretching vibration located around 1083 cm^{-1} in the spectra cannot be used as an internal standard for spectral normalization as has been done by other workers (Akhebat et al., 1992). This is because association of 23-mer and 7-mer RNA affects this band. Also in case of 7-mer RNA this band is unusually intense and the intensity is strongly correlated with the intensity of the 945 cm^{-1} band when the temperature or Mg^{2+} concentration was changed. However, a complete overlap of the antisymmetric PO_2^- stretching vibration at 1241 cm^{-1} between the added spectrum of the RNA component strands and the complex was observed. Also in case of the DNA strands where position of the 1083 cm^{-1} band was not affected in the 1:1 mixture, a complete overlap was observed between the added and the spectrum of the mixture. This shows that it is sufficient to compare the added spectrum of the components with that of the complex without additional normalization of the FTIR spectra of individual RNA/DNA strands.

CONCLUSION

FTIR-monitored association of the two strand model system of the P4−P6 triple-helical domain of the group I intron core provides evidence for the formation of the P6 helix in the complex, correctly modeling this domain. Our results show that the proper assembly of the triple-helical domain requires a ribose backbone. Changes in the backbone−sugar frequency region (1500−750 cm^{-1}) show that the interactions between the two strands involve the riboses which in turn could also involve the 2′-OH groups. Studies on 7-mer RNA variants with single deoxyribose substitution at specific positions show that either the absence of 2′-OH groups at these positions and/or changes in the local structure perturb the assembly of the triple-helical domain. FTIR-monitored melting curves show ordering of the non-paired

end of the stem duplex in 23-mer RNA, but the deoxyribose backbone of the 23-mer DNA hairpin was incapable of such ordering.

REFERENCES

Akhebat, A., Dagneaux, C., Liquier, J., & Taillandier, E. (1992) *J. Biomol. Struct. Dyn. 10*, 577−588.

Cate, J. H., Gooding, A. R., Podell, E., Zhou, K., Golden, B. L., Kundrot, C. E., Cech, T. R., & Doudna, J. A. (1996) *Science 273*, 1678−1685.

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

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

Connolly, B. A. (1991) in *Oligonucleotides and Analogues: A Practical Approach* (Eckstein, F., Ed.) pp 161−162, IRL Press, Oxford.

Dagneaux, C., Liquier, J., Scaria, P. V., Shafer, R. H., & Taillandier, E. (1994) in *Structural Biology: The State of the Art, Proc. of the VIIIth Conversation, State University of New York, Albany, New York, 1993* (Sarma, R. H., & Sarma, M. H., Eds.) pp 103−111, Adenine Press, Schenectady, NY.

Dagneaux, C., Liquier, J., & Taillandier, E. (1995) *Biochemistry 34*, 14815−14818.

Dohy, D., Ghomi, M., & Taillandier, E. (1989) *J. Biomol. Struct. Dyn. 6*, 741−754.

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

Holland, J. A., & Hoffman., D. W. (1996) *Nucleic Acids Res. 24*, 2841−2848.

Klinck, R., Guittet, E., Liquier, J., Taillandier, E., Gouyette, C., & Huynh-Dinh, T. (1994) *FEBS Lett. 355*, 297−300.

Letellier, R., Ghomi, M., & Taillandier, E. (1989) *J. Biomol. Struct. Dyn. 6*, 755−768.

Li, S.-J., Burkey, K. O., Luoma, G. A., Alben, J. O., & Marshall, A. G. (1984) *Biochemistry 23* 3652−3658.

Liquier, J., Akhebat, A., Taillandier, E., Ceolin, F., Huynh-Dinh, T., & Igolen, J. (1991) *Spectrochim. Acta A 47*, 177−186.

Liquier, J., Taillandier, E., Klinck. R., Guittet, E., Gouyette, C., & Huynh-Dinh, T. (1995) *Nucleic Acids Res. 23*, 1722−1728.

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

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.

Nowakowski, J., & Tinoco, I., Jr. (1996) *Biochemistry 35*, 2577−2585.

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

Quali, M., Letellier, R., Sim, J.-S., Akhebat, A., Adnet, F., Liquier, J., & Taillandier, E. (1993) *J. Am. Chem. Soc. 115*, 4264−4270.

Sarkar, M., Sigurdsson, S., Tomac, S., Sen. S., Rozners, E., Sjöberg, B.-M., Strömberg, R., & Gräslund A. (1996) *Biochemistry 35*, 4678−4688.

Schulhof, J. C., Molko, D., & Teoule, R. (1987) *Nucleic Acids Res. 15*, 397−416.

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

Taboury, J. A., Liquier, J., & Taillandier, E. (1985) *Can. J. Chem. 63*, 1904−1909.

Taillandier, E., & Liquier, J. (1992) *Methods Enzymol. 211*, 307−335.

Taillandier, E., Liquier, J., & Taboury, J. A. (1985) in *Advances in Infrared and Raman Spectroscopy* (Clark, R. H., & Hester, R. E., Eds.) Vol. 12, pp 65−114, Wiley-Heyden, New York.

Tsuboi, M. (1969) in *Applied Spectroscopy Reviews* (Brame, E. G., Jr., Ed.) Vol. 3, pp 45−90, Dekker, New York.

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

BI9702243