Melançon, P., Lemieux, C., & Brakier-Gingras, L. (1988) *Nucleic Acids Res. 16*, 9631–9639.
Moazed, D., & Noller, H. F. (1986) *Cell 47*, 985–994.
Moazed, D., & Noller, H. F. (1990) *J. Mol. Biol. 211*, 135–145.
Moazed, D., Robertson, J. M., & Noller, H. F. (1988) *Nature 334*, 362–364.
Muralikrishna, P., & Wickstrom, E. (1989) *Biochemistry 28*, 7505–7510.
Noller, H. F., Kop, J. A., Wheaton, V., Brosius, J., Gutell, R. R., Kopylov, A. M., Dohme, F., Stahl, D. A., Gupta, R., & Woese, C. R. (1981) *Nucleic Acids Res. 9*, 6167–6189.
Oakes, M. I., Kahan, L., & Lake, J. A. (1990) *J. Mol. Biol. 211*, 907–918.
Ofengand, J., Ciesiolka, J., Denman, R., & Nurse, K. (1986) in *Structure, Function, and Genetics of Ribosomes* (Hardesty, B., & Kramer, G., Eds.) pp 473–494, Springer-Verlag, New York.
Ryan, P. C., & Draper, D. E. (1989) *Biochemistry 28*, 9949–9956.
Scheit, K. H. (1968) *Chem. Ber. 101*, 1147.
Sköld, S. E. (1983) *Nucleic Acids Res. 11*, 4923–4932.
Spirin, A. S., & Vasiliev, V. D. (1989) *Biol. Chem. 66*, 215–223.
Stade, K., Rinke-Appel, J., & Brimacombe, R. (1989) *Nucleic Acids Res. 17*, 9889–9908.
Staehelin, T., & Maglott, C. R. (1971) *Methods Enzymol. 20*, 449–456.
Stern, S., Weiser, B., & Noller, H. F. (1988) *J. Mol. Biol. 204*, 447–481.
Stiege, W., Stade, D., Schuler, D., & Brimacombe, R. (1988) *Nucleic Acids Res. 16*, 2369–2388.
Stoeffler, G., & Stoeffler-Meilicke, M. (1986) in *Structure, Function, and Genetics of Ribosomes* (Hardesty, B., & Kramer, G., Eds.) pp 28–46, Springer-Verlag, New York.
Tanner, N. K., Hanna, M. M., & Abelson, J. (1988) *Biochemistry 27*, 8852–8861.
Thompson, J., Schmidt, F., & Cundliffe, E. (1982) *J. Biol. Chem. 257*, 7915–7917.
Traub, P., Mizushima, S., Lowry, C. V., & Nomura, M. (1971) *Methods Enzymol. 20*, 391–407.
Trempe, M. R., Ohgi, K., & Glitz, D. G. (1982) *J. Biol. Chem. 257*, 9822–9829.
Trifonov, E. N. (1987) *J. Mol. Biol. 194*, 643–652.
Van Duin, K., Kurland, C. G., Dondon, J., & Grunberg-Manago, M. (1975) *FEBS Lett. 59*, 287–290.
Vladimirov, S. N., Babkina, G. T., Venijaminova, A. G., Gimautdinova, O. I., Zenkova, M. A., & Karpova, G. G. (1990) *Biochim. Biophys. Acta 1048*, 245–256.
Wagner, R., Gassen, H. G., Ehresmann, Ch., Stiegler, P., & Ebel, J. P. (1976) *FEBS Lett. 67*, 312–315.
Walseth, T. F., & Johnson, R. A. (1979) *Biochim. Biophys. Acta 562*, 11–31.
Weiss, R. B., Dunn, D. M., Dahlberg, A. E., Atkins, J. F., & Gesteland, R. F. (1988) *EMBO J. 7*, 1503–1507.
Wollenzien, P. L. (1988) *Methods Enzymol. 164*, 319–329.

Properties of a U1/mRNA 5′ Splice Site Duplex Containing Pseudouridine As Measured by Thermodynamic and NMR Methods†

Kathleen B. Hall*
Department of Biochemistry and Molecular Biophysics, Washington University School of Medicine, Box 8231, 660 South Euclid Avenue, St. Louis, Missouri 63110
Larry W. McLaughlin*
Department of Chemistry, Boston College, 140 Commonwealth Avenue, Chestnut Hill, Massachusetts 02167
Received September 5, 1990; Revised Manuscript Received November 27, 1990

ABSTRACT: Three RNA undecamers, $^{5'}$AUACΨΨACCUG (Ψ = pseudouridine), $^{5'}$AUACUUACCUG, and their complementary 11-mer $^{5'}$CAGGUAAGUAU, have been chemically synthesized by phosphite triester chemistry on a controlled-pore glass (CPG) support. The two duplexes formed with these molecules, $^{5'}$AUACΨΨACCUG/$^{5'}$CAGGUAAGUAU and $^{5'}$AUACUUACCUG/$^{5'}$CAGGUAAGUAU, represent the 5′ end of human U1 snRNA paired to the mRNA consensus 5′ splice site. In one undecamer, pseudouridines are incorporated at those positions corresponding to the native in vivo U1 snRNA, while the other (control) undecamer contains only uridine. Surprisingly, the NMR data show that the extra imino proton of the pseudouridines, which is found in the major groove and is presumably not hydrogen bonded, is clearly visible in the imino proton NMR spectrum at pH 6. This result suggests that the structure of the RNA restricts access of solvent to the major groove, slowing the exchange of the pseudouridine NH1 imino proton. A comparison of the thermodynamic properties of the two duplexes show that the free energy of duplex formation is unchanged by the substitution of pseudouridine for uridine.

Most RNA molecules contain modified nucleotide residues at specific locations within a given sequence. The functional and/or structural importance of these modifications is not well understood.

Pseudouridine (Ψ) is a regioisomer of uridine in which the heterocyclic base uracil is bound to the ribose sugar through the carbon atom at the C5 position rather than through the normal N1 nitrogen (Figure 1). With the formation of a C-nucleotide in place of the more common N-nucleotide, two

†Supported by the Lucille P. Markey Charitable Trust (88-30 to K.B.H.) and by grants to L.W.M. from the National Science Foundation (DMB-8904306) and the National Institutes of Health (GM-37065).
*Correspondence can be addressed to either author.

0006-2960/91/0430-1795$02.50/0

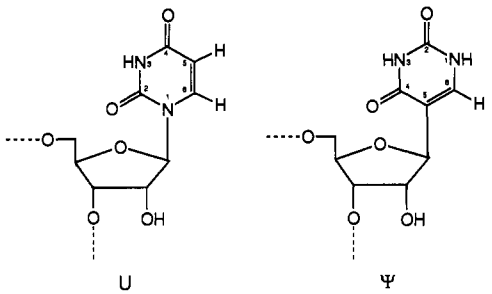

FIGURE 1: Structure of uridine and pseudouridine.

imino hydrogens rather than one are present in pseudouridine. In double-stranded RNA, the N3-imino hydrogen can be expected to form a Watson–Crick hydrogen bond with adenine. The extra N–H functionality, located in the major groove of the duplex, may influence RNA structure by providing an additional or alternate hydrogen-bonding site, or possibly by providing a contact to stabilize protein–RNA complexes. Pseudouridine is commonly found in the sequence of tRNAs (Sprinzl et al., 1989) as well as that of the small nuclear RNAs (snRNAs). It is present in both single- and double-stranded regions.

In human U1 snRNA, two pseudouridine residues are present in the 5′ end of human U1 snRNA (Reddy, 1989). In the proposed structure of U1 snRNA, the 5′-terminal 12 nucleotide residues including the cap structure and the two pseudouridines are apparently single stranded (Krol et al., 1990) and thus available to pair with the mRNA 5′ splice site (Zhuang & Weiner, 1986; Seraphin et al., 1988; Siliciano & Guthrie, 1988), contributing to splice site recognition. Many species, from human to *Drosophila*, include these two pseudouridines at this position in their U1 snRNA (Reddy, 1989), which suggests that pseudouridine may have an important structural/functional role. In addition to pseudouridines, the 5′ end of native human U1 snRNA contains a trimethylguanosine cap, as well as 2′-*O*-methyluridine and 2′-*O*-methyladenosine residues.

In this work, two undecamers corresponding to the 5′-terminal single-stranded human U1 snRNA sequence (excluding the cap structure) were chemically synthesized: the "native" sequence contains two pseudouridines at positions 5 and 6 ($^{5'}A_1UAC\Psi\Psi ACCUG_{11}$) while the "control" sequence ($^{5'}$AUACUUACCUG) contains uridines. The complementary undecamer ($^{5'}C_{12}AGGUAAGUAU_{22}$) was designed to contain the consensus mRNA 5′ splice site (Mount, 1982). The two duplexes formed from these undecamers allow a comparative study of the structural and thermodynamic properties of pseudouridine in double-stranded RNA and possibly some insight into metazoan splice site recognition.

Materials and Methods

Materials. The individual nucleoside phosphoramidites were purchased from Milligen-Biosearch (Burlington, MA), and the controlled-pore glass support was from Peninsula Laboratories, Inc. (Belmont, CA). Pseudouridine in the protected phosphoramidite form was synthesized according to the procedures of Ogilvie (Hakimelahi et al., 1982) and Narang (Sung & Narang, 1982) and will be described elsewhere.

Oligonucleotide Synthesis. $^{5'}$AUACUUACCUG, $^{5'}$AUAC$\Psi\Psi$ACCUG, and $^{5'}$CAGGUAAGUAU were chemically synthesized on controlled-pore glass supports by phosphite triester chemistry (Usman et al., 1987). The synthesis employed 10 μmol of bound nucleoside in a glass column. Solvents and reagents were delivered to the column by use of an Applied Biosystems 380A DNA synthesizer with the argon pressure set at 4.2 bar. After deprotection (tetrabutylammonium fluoride, 16 h, ambient temperature) of the assembled oligonucleotides (Usman et al., 1987), the products were purified by FPLC on a 1.0 × 10 cm MonoQ column at a flow rate of 4 mL/min in 5 mM sodium cacodylate, pH 6.0, and with a gradient of sodium chloride (0–0.33 M in 25 mL followed by 0.33–0.48 M in 90 mL). The undecamers typically eluted in the range of 67–74 mL. After isolation and desalting (Sephadex G-10), the products were analyzed by HPLC on a 0.46 × 25 cm column of ODS-Hypersil at a flow rate of 1.5 mL/min in 20 mM sodium phosphate, pH 5.5, with a linear gradient of methanol (0–70% in 60 min). If necessary, a second purification was undertaken by HPLC on a 0.94 × 25 cm column of ODS-Hypersil and with the gradient described above.

Nucleotide (or nucleoside) composition was determined after S1 nuclease (or S1 nuclease and bacterial alkaline phosphatase) hydrolysis. A 10-μL reaction mixture containing 1 A_{260} unit of undecamer in 200 mM NaCl/5 mM $MgCl_2$/0.1 mM $ZnSO_4$/25 mM sodium acetate, pH 5.5, was incubated for 5 min at room temperature with 267 units of S1 nuclease. A 3-μL aliquot was analyzed by HPLC, with a 60-min 0–35% methanol gradient. For nucleoside analysis, 5 μL of 0.1 M Tris-HCl, pH 8.0, and 1 unit of bacterial alkaline phosphatase were added to the remaining 7 μL of digested undecamer. Following incubation for 60 min at ambient temperature, a 5-μL aliquot was analyzed by HPLC on a 60-min 0–70% methanol gradient.

Melting Curves. Extinction coefficients for each strand were calculated with the nearest-neighbor approximation (Richards, 1975). Values at 260 nm ($\times 10^{-4}$ cm^{-1} M^{-1}) are 112 for AUACUUACCUG and AUAC$\Psi\Psi$ACCUG and 116 for CAGGUAAGUAU. In the absence of data for pseudouridine, the corresponding values for uridine were used in these calculations. The extinction coefficient of free uridine at 260 nm, pH 7, is 10.1 ($\times 10^{-4}$ cm^{-1} M^{-1}), and that of pseudouridine is 8.1 (Shapiro & Chambers, 1961; Chambers, 1966), so there will be some error introduced by this method. Strand concentrations were determined from the absorbance measured at 90 °C.

Melting curves of absorbance vs temperature were recorded on the Gilford 250 system in the laboratory of Prof. D. Turner (Department of Chemistry, University of Rochester) as described previously (Freier et al., 1983). Each single strand was measured separately at two concentrations; the melting temperatures were obtained for nine concentrations of pseudouridine duplexes and for six concentrations of uridine duplexes, spanning a 100-fold concentration range (1×10^{-4} to 1×10^{-6} M strands). The buffer used was 1.0 M NaCl, 10 mM sodium cacodylate, and 0.5 mM EDTA, pH 7. Heating rate for the melting experiments was 1 °C/min.

Data were analyzed as described previously (Petersheim & Turner, 1983; Freier et al., 1983). Two methods were used to calculate thermodynamic parameters of helix formation: data were plotted as log C_T vs $1/T_M$ to give enthalpy (ΔH) and entropy (ΔS); each melting curve was fit to a two-state model to give enthalpy and entropy values which were then averaged.

NMR. Each single strand was dissolved in 0.1 M NaCl. The oligonucleotide concentrations were determined from the absorbance at 260 nm of 1/2000 dilutions at 30 °C, and the strands were then mixed proportionately to make the duplexes. The addition of 1 M sodium cacodylate, pH 6, 0.1 M NaCl, and D_2O gave final concentrations of 10 mM sodium cacodylate, 0.1 M NaCl, 10% D_2O, and 2 mM duplex in a total

of 350 μL. Before spectra were recorded, the samples were heated to 65 °C and cooled slowly.

Imino proton NMR spectra were recorded by use of a 214 pulse for water suppression (Redfield et al., 1975) with the carrier centered in the imino proton region. The spectral width was 8 kHz, and the recovery time was 600 ms. NOE difference spectra were recorded in the interleaved mode with the off-resonance spectrum. Spectra were recorded on the 500-MHz spectrometer in Prof. A. G. Redfield's laboratory at Brandeis University.

Results

Oligonucleotide Synthesis. The three RNA undecamers were prepared by solid-phase phosphite triester synthesis on a controlled-pore glass support. Phosphoramidite building blocks which contained a 2′-*O*-(*tert*-butyldimethylsilyl) residue, β-cyanoethoxy phosphorus, and isobutyryl (G) or benzoyl (A, C) base protecting groups were used essentially as described by Usman et al. (1987). Incorporation of the pseudouridine phosphoramidite into the growing oligonucleotide chain occurred with yields comparable to those obtained with the "normal" monomers. From an initial 10 μmol of 3′-bound nucleoside, we were able to purify approximately 10–12 mg (200–240 A_{260} units) of each strand. Purification proceeded primarily by anion-exchange chromatography, but a second isolation using reversed-phase chromatography was required in some cases.

Nuclease digestion using S1 nuclease confirmed the fidelity of the 3′–5′ phosphodiester linkages (Figure 2a). The regiochemistry at the phosphodiester linkage is a necessary concern since the preparation of the nucleoside building blocks requires the separation of the 2′-silyl and 3′-silyl derivatives (Ogilvie et al., 1979). If some contamination by the 3′-silyl derivatives remains after this step, 2′–5′ phosphodiester linkages will be formed in place of the normal 3′–5′ linkages. The 2′–5′ phosphodiesters are not substrates for S1 nuclease; thus, the presence of such linkages would result in incomplete digestion of the fragments with residual nucleoside 2′–5′ dimers present. The digestion products were resolved by HPLC, and even with extreme overloading of the HPLC column, no evidence of such dimers was found. Digestion by S1 nuclease produced 5′-nucleotide monophosphates with the exception of the 5′-terminal residue which eluted as a nucleoside (Figure 2a).

Nucleotide/nucleoside composition for the unmodified sequences was determined by analysis of the products of the S1 nuclease experiments (Figure 2a). However, pseudouridine 5′-monophosphate could not be effectively resolved from cytidine 5′-monophosphate under the described conditions. In this case, S1 nuclease and bacterial alkaline phosphatase were used to digest the Ψ-containing fragment to nucleosides. HPLC analysis was then used to confirm the presence of two Ψ residues (Figure 2b).

It has been reported (Cohn, 1960; Chambers et al., 1963) that upon treatment of pseudouridine with 1 N HCl at 100 °C for several hours or under strong alkaline conditions some isomerization of the carbohydrate residue occurred to produce Ψ_B, Ψ_{AF}, and Ψ_{AS} isomers, which are distinguished chromatographically. Prior to the synthesis of the pseudouridine-containing undecamer, we treated the free nucleoside with concentrated ammonia at 50 °C for 18 h. HPLC analysis of this mixture indicated that only the β-anomer was present and no isomerization had occurred (data not shown). In the present oligonucleotide synthesis, each elongation cycle during assembly of the RNA molecule begins with treatment by acid, but under milder conditions (trichloroacetic acid for 2 min at ambient temperature). Under these conditions, and after 6-h incubation in ammonia, we were unable to detect any pseudouridine isomer in the nucleoside analysis of the product (Figure 2b) other than the natural 5′-β-D-ribofuranosyluracil.

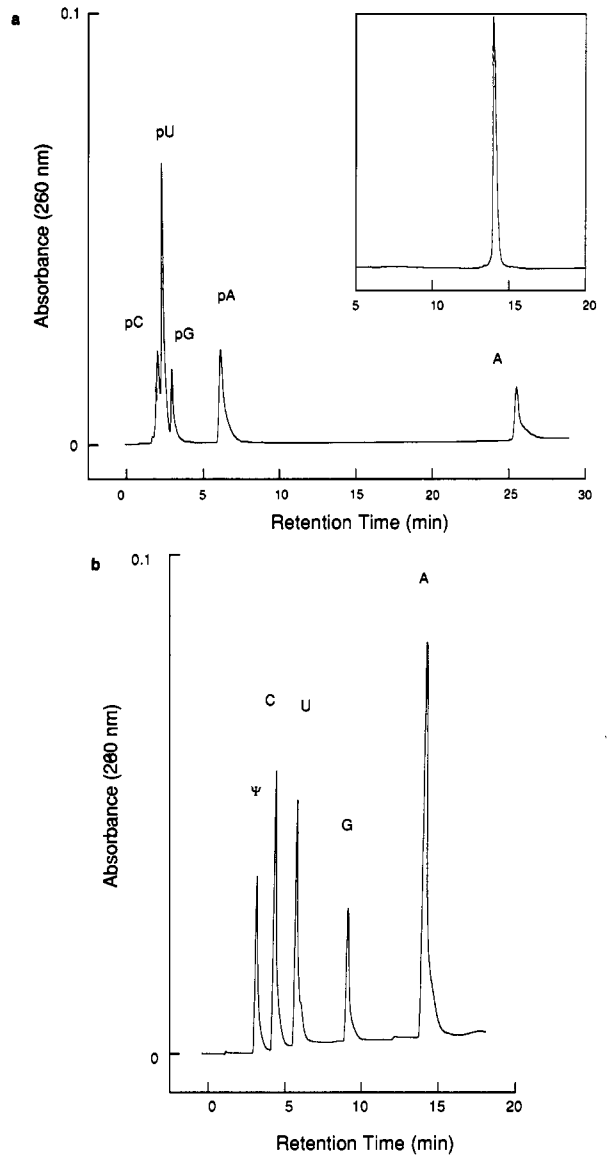

FIGURE 2: (a) Reversed-phase HPLC analysis of the products resulting from S1 nuclease digestion of the undecamer 5′AUACUUACCUG under the following chromatographic conditions: 0.46 × 25 cm ODS-Hypersil column; flow rate 1.5 mL/min; 20 mM KH_2PO_4, pH 5.5; 60-min 0–35% methanol gradient. Inset: analysis of the intact undecamer by reversed-phase chromatography under the above conditions with a gradient of 0–70% methanol in 60 min. (b) Reversed-phase HPLC analysis of the products resulting from S1 nuclease and bacterial alkaline phosphatase digestion of the undecamer 5′AUACΨΨACCUG. Conditions as above with a 30-min 0–35% methanol gradient.

Thermodynamic Parameters. Plots of log (concentration) vs the reciprocal melting temperatures ($1/T_M$) of the two duplexes are shown in Figure 3. The calculated entropy (ΔS) and enthalpy (ΔH) from the plots of Figure 3 are given in Table I. Also given in Table I are those values derived from fitting the melting curves to the two-state approximation. The values from the two methods agree within experimental error, as would be expected if the duplexes underwent a two-state helix–coil transition. Table I includes the calculated values of enthalpy and entropy from the nearest-neighbor model (Borer et al., 1974) and the values from Freier et al. (1986) and Turner et al. (1988).

Table I: Temperature-Independent Thermodynamic Parameters[a]

duplex	av from fits[b]			$1/T_M$ vs log C_T[b]			calcd[c]		
	$-\Delta H°$	$-\Delta S°$	T_M[d]	$-\Delta H°$	$-\Delta S°$	T_M[d]	$-\Delta H°$	$-\Delta S°$	T_M[d]
AUACUUACCUG UAUGAAUGGAC	96.5	264	64.7	95.3	261	64.7	86.8	238	61
AUACΨΨACCUG UAUGAAUGGAC	92.5	250	67.3	94.3	256	67.2			

[a] Buffer conditions: 1 M NaCl, 10 mM sodium cacodylate, 0.5 mM EDTA, and pH 7. $\Delta H°$ is in units of kcal/mol, $\Delta S°$ in eu, and T_M in °C. Errors in $\Delta S°$ and $\Delta H°$ are ±5%. [b] Methods as described in Petersheim and Turner (1983) and Freier et al. (1983). For noncomplementary strands, the calculations are for $C_T/4$. [c] Calculated as per Turner et al. (1988). [d] Calculated for 1×10^{-4} M strands.

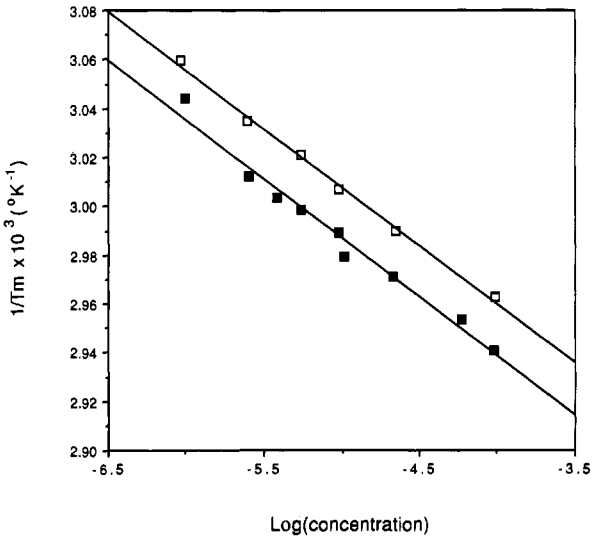

FIGURE 3: Reciprocal melting temperature vs log (concentration) for the two duplexes. Open squares, U duplex; solid squares, Ψ duplex.

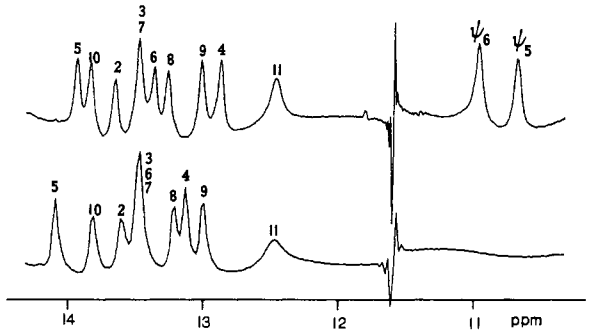

FIGURE 4: Imino proton spectra of Ψ duplex (top) and U duplex (bottom) at 25 °C. Assignments are indicated, following the numbering of the U strand: $A_1U_2A_3C_4U(\Psi)_5U(\Psi)_6A_7C_8C_9U_{10}G_{11}$. The signal at 11.6 ppm is a machine artifact at the carrier. Chemical shifts are relative to water at 4.8 ppm.

Imino Proton Spectra. Imino proton NMR spectra of the two duplexes at 25 °C are shown in Figure 4. While there are potentially 11 base pairs formed in the duplex, only 10 imino protons are found in the region between 14.5 and 12.0 ppm. In the spectrum of the pseudouridine duplex, there are two additional imino protons, at 10.92 and 10.65 ppm.

The imino proton spectra of the uridine duplex are shown as a function of temperature in Figure 5. Six resonances remain at 55 °C, which disappear together at 65 °C (data not shown). Figure 6 shows analogous data for the pseudouridine duplex. Again, all resonances disappear at 65 °C.

Imino Proton Assignments. The AU base pairs were assigned by use of the strong intrabase NOE from the imino proton to the adenosine H2 resonance. Counting the number of AU base pairs identified shows that the missing resonance is from an AU base pair, assigned to A_1U_{22}.

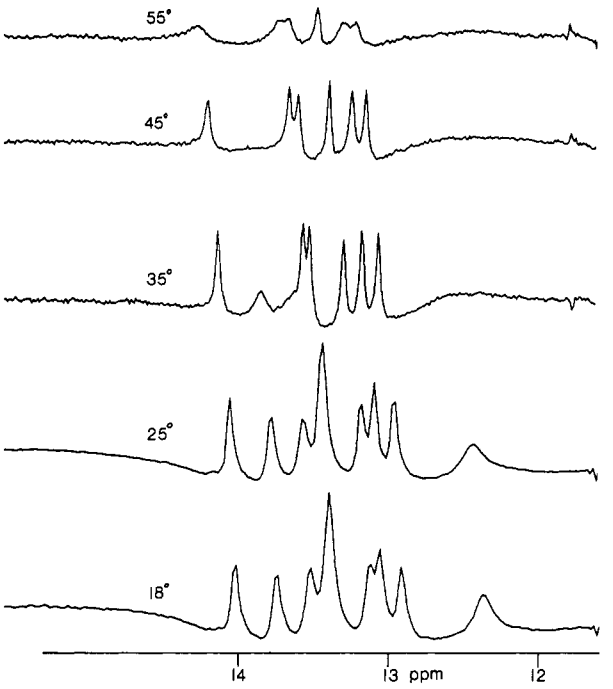

FIGURE 5: Temperature dependence of the imino proton spectrum of the U-containing duplex.

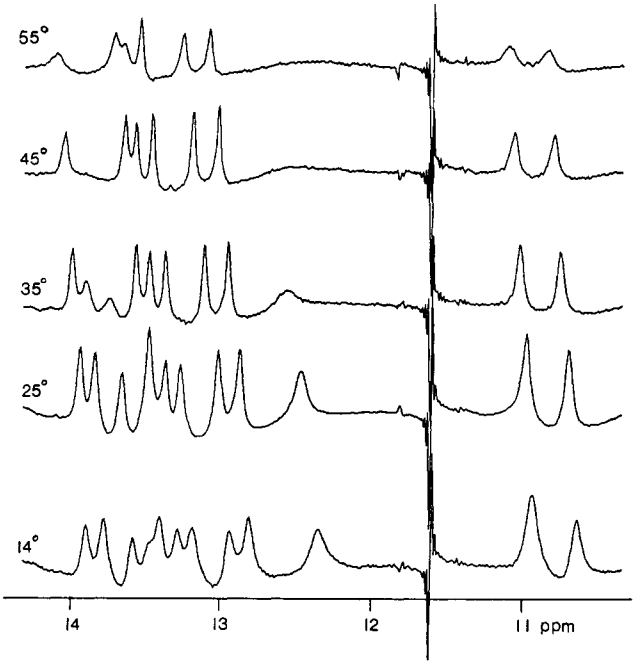

FIGURE 6: Temperature dependence of the imino proton spectrum of the Ψ-containing duplex. The signal at 11.6 ppm is a machine artifact at the carrier.

Using inter-base-pair NOEs between imino protons, assignments were made for the individual resonances, and these

Table II: Imino Proton Assignments[a]

base pair	U duplex (ppm)[b]	Ψ duplex (ppm)[b]	difference
A_1U_{22}	not found		
U_2A_{21}	13.57	13.61	0
A_3U_{20}	13.45	13.44	0
C_4G_{19}	13.09	12.81	0.28
$U_5(\Psi)A_{18}$	14.06	13.9	0.16
$U_6(\Psi)A_{17}$	13.45	13.32	0.13
A_7U_{16}	13.45	13.45	0
C_8G_{15}	13.14	13.22	−0.08
C_9G_{14}	12.96	12.98	0
$U_{10}A_{13}$	13.77	13.8	0
$G_{11}C_{12}$	12.47	12.45	0
Ψ_5		10.64	
Ψ_6		10.93	

[a] Measured at 25 °C in 0.1 M NaCl–10 mM sodium cacodylate, pH 6. [b] Errors in chemical shifts are ±0.04 ppm.

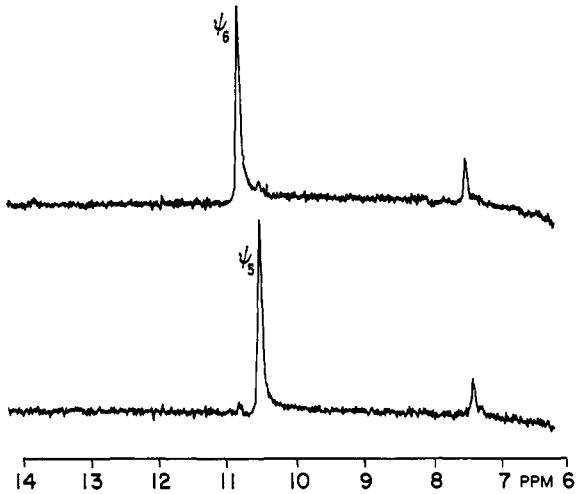

FIGURE 7: NOE data from the extra pseudouridine imino protons. Irradiation time was 80 ms and recovery time 400 ms, at 20 °C. The strong NOEs to the aromatic region, at 7.55 ppm (top) and 7.45 ppm (bottom), identify the adjacent CH6 protons.

are given in Table II. No NOEs were observed to or from the resonance at 12.45 ppm, leading to its assignment as the terminal GC pair. The broad line width of this resonance suggests that the base pair is fraying.

The two additional imino protons in the spectrum of the pseudouridine duplex show NOEs to each other as well as to resonances in the aromatic region. These data are shown in Figure 7. Those imino resonances are thus assigned to the N1H imino proton and the corresponding aromatic resonances to the adjacent CH6 proton on each pseudouridine base. The imino to imino NOE is analogous to the NOE from adjacent pyrimidine H5/pyrimidine H5 nonexchangeable protons seen in A-form structures where these protons are 3.7 Å apart (compared to 4.7 Å apart in B-form) (Wuthrich, 1986).

Discussion

These experiments have provided new information on the nucleic acid physical chemistry of pseudouridine incorporation into RNA helices as well as some insight into the interaction between the 5′ end of U1 snRNA and the mRNA 5′ splice site.

Pseudouridine and Duplex Structure. The NMR imino proton spectrum of the pseudouridine duplex is much more disperse than that of the uridine duplex. Table II gives the chemical shift difference between analogous imino protons in the two duplexes. The largest difference is seen not in the AΨ vs AU pairs but in C_4G_{19} on the 5′ side of Ψ_5A_{18}. The 0.28 ppm upfield shift of the C_4G_{19} imino proton in the pseudouridine duplex suggests that there is additional shielding from adjacent ring systems, due to more overlap in the stacking, to ring current effects, or to both effects. Only this GC and the two AΨ pairs show significantly different imino proton chemical shifts in the two duplexes, indicating that any conformational change is local. More detailed NMR information, using the nonexchangeable protons, will provide information about possible structural alteration of the duplex as a result of the pseudouridine residues.

Extra Imino Protons. Imino protons are exchangeable and therefore are observed most often in NMR (H_2O) spectra only when they are hydrogen bonded or protected from solvent. An upfield chemical shift of an imino proton, when the resonances are found between 10 and 11.5 ppm, is a signature of a non-hydrogen-bonded proton (Haasnoot et al., 1980, 1983). Such protons from both RNA and DNA are usually found in loops (Haasnoot et al., 1983, 1986; Ikuta et al., 1986; Puglisi et al., 1990) or in bulges (Varani et al., 1989) where the structure impedes solvent accessibility. However, one of the hydrogen-bonded imino protons of GU pairs is also found between 10 and 11.5 ppm (Johnston & Redfield, 1978, 1981). Thus, chemical shift alone is insufficient to determine the hydrogen-bonding state of an imino proton.

Resonances from the pseudouridine N1 imino protons are observed in the NMR spectrum of exchangeable protons. This imino proton resonance was observed previously in the NMR spectrum of the yeast $tRNA^{Phe}$ anticodon stem and loop (Clore et al., 1984) where the AΨ base pair is contiguous to the loop. Both the crystal structure of yeast $tRNA^{Phe}$ (Holbrook et al., 1978) and that NMR study of the anticodon stem/loop showed that this loop is highly structured and therefore could protect the adjacent pseudouridine proton from solvent.

Our results indicate that the formation of the A-form duplex impedes exchange of the pseudouridine N1 imino proton. The N1 imino proton of pseudouridine is found in the major groove of the RNA duplex. Its stability could be due to hydrogen bonding (possibly to a phosphate oxygen) or to protection from exchange by the major groove structure. Alternatively, its stability could reflect the intrinsic exchange property of this proton (Fritzsche et al., 1981: McConnell & Seawall, 1972), although preliminary data suggest that, under the solution conditions of these experiments, the exchange rate of this proton in the pseudouridine nucleoside is too fast to measure (data not shown). Pochon et al. (1964) reported that the pK of the N1 proton of pseudouridine was identical in the monomer and in poly(Ψ) (pK = 9.7). The pH dependence of the properties of these NH1 resonances will provide additional data on the cause of their stability.

The thermal stability of the N1 imino proton resonances is surprising and may provide a clue to the source of its exchange properties. As shown in Figure 5, the temperature dependence of the intensity of these proton resonances follows that of the base-paired imino protons and is thus similar to the behavior of thymidine loop protons in DNA (Haasnoot et al., 1983). In the experiments of Haasnoot et al., the loss of peak intensity from those non-hydrogen-bonded thymidine NH3 loop protons is correlated to the disruption of the structure of both the loop and the duplex, such that the imino protons become accessible to the bulk water and thereby increase their exchange rate. In our pseudouridine experiments, the loss of NH1 proton resonance intensity follows that of the hydrogen-bonded NH3 protons. Normally, exchange of the NH3 protons increases as the duplex dissociates or the base pairs fray and is manifested by the temperature-dependent

Table III: Free Energy of Duplex Formation[a]

duplex	av from fits[b]			$1/T_M$ vs log C_T[b]			calcd
	$-\Delta G°_{50}$	$-\Delta G°_{37}$	$-\Delta G°_{25}$	$-\Delta G°_{50}$	$-\Delta G°_{37}$	$-\Delta G°_{25}$	$-\Delta G°_{37}$[c]
AUACUUACCUG UAUGAAUGGAC	11	14.4	17.6	11	14.4	17.5	12.9
AUACΨΨACCUG UAUGAAUGGAC	11.5	14.8	17.8	11.6	14.9	18	

[a] Buffer conditions: 1 M NaCl, 10 mM sodium cacodylate, 0.5 mM EDTA, and pH 7. [b] Values for $\Delta G°$ are in kcal/mol. Errors are ±5%. [c] Calculated as per Turner et al. (1988).

disappearance of the imino proton resonances. The pseudouridine NH1 proton resonances disappear at the same temperature as hydrogen-bonded imino proton resonances from the core hexamer, suggesting that only by disruption of the duplex is solvent water accessible to the NH1 protons in the major groove. However, the line widths of the NH1 imino protons do broaden more rapidly than those of the base-paired protons, especially above 40 °C when only the six core base pairs are present, suggesting that the NH1 protons can be more readily exchanged than corresponding Watson–Crick hydrogen-bonded protons.

An exchangeable imino proton in the major groove provides a unique NMR probe of the environment and dynamics of the A-form helix. It can potentially provide details of base-pair opening, as well as report conformational changes that result from ligand binding or protein interactions.

Thermodynamics of Pseudouridine Incorporation. The data of Table I are used to calculate the free energy of duplex formation for the two duplexes. The ΔG values determined from the different methods are given in Table III and are compared to the value from the nearest-neighbor model. Clearly, the presence of pseudouridine does not alter helix structure enough to change the free energy of the duplex.

One hypothesis to explain the presence of pseudouridine in U1 snRNA is that the modified base stabilized the U1/mRNA duplex. Previous experiments (Pochon et al., 1964) had shown that poly(A)/poly(Ψ) duplexes had higher melting temperatures than poly(A)/poly(U) duplexes. Any stabilization would be especially important when the U1/mRNA duplex included unpaired bases, as would be the case for many in vivo splice sites (Mount, 1982). The results observed here, where pseudouridine duplexes show only a 2 °C increase in the T_M, indicate that such stabilization cannot be expected; thus, another function for this base must be sought. It is still possible that, in the case where the pseudouridines are adjacent to an unpaired nucleotide, they could provide some stabilization through stacking interactions or hydrogen bonding.

The temperature dependence of the NMR imino proton spectrum identifies a core of six base pairs that is more stable than the flanking sequences. This core sequence, whether $^{5'}$CUUACC/$^{5'}$GGUAAG or $^{5'}$CΨΨACC/$^{5'}$GGUAAG, appears to melt cooperatively at 65 °C. In vivo, this core duplex would contain the paired mRNA 5′ splice site, $G_{14}G_{15}U_{16}$, (Zhuang and Weiner, 1986) in which the GU dimer sequence of the intron is absolutely conserved (Mount, 1982).

Summary

We have shown that the additional imino proton of pseudouridine, because it is exchangeable, is a new NMR probe of the major groove of RNA. It can be used both for distance measurements and for potential observation of both base-pair opening and the solution environment of the A-form major groove. On the basis of preliminary spectral data, pseudouridine does not appear to be an isomorphic replacement for uridine, but the extent of the difference cannot yet be determined. Incorporation of pseudouridine into an RNA duplex neither destabilizes nor stabilizes formation of the double strand. The intron/exon junction is found in a stable hexamer core of the both U and Ψ duplexes.

Acknowledgments

We thank Doug Turner for his generosity and hospitality, John SantaLucia for his time and patience, Adam Peritz for his help with the computer, and Al Redfield for the time on his spectrometer. K.B.H. thanks Art Pardi for very helpful discussions.

References

Borer, P. N., Dengler, B., Tinoco, I., Jr., & Uhlenbeck, O. C. (1974) *J. Mol. Biol. 86*, 843–853.

Chambers, R. W. (1966) *Prog. Nucleic Acid Res. Mol. Biol. 5*, 349–398.

Chambers, R. W., Kurkov, V., & Shapiro, R. (1963) *Biochemistry 2*, 1192–1203.

Clore, G. M., Gronenborn, A. M., Piper, E. A., McLaughlin, L. W., Graeser, E., & Van Boom, J. H. (1984) *Biochem. J. 221*, 737–751.

Cohn, W. E. (1960) *J. Biol. Chem. 235*, 1488–1498.

Freier, S. M., Burger, B. J., Alkema, D., Neilson, T., & Turner, D. H. (1983) *Biochemistry 22*, 6198–6206.

Freier, S. M., Kierzek, R., Jaeger, J. A., Sugimoto, N., Caruthers, M. H., Neilson, T., & Turner, D. H. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 9373–9377.

Fritzsche, H., Kan, L.-S., & Ts'o, P. O. P. (1981) *Biochemistry 20*, 6118–6122.

Haasnoot, C. A. G., den Hartog, J. H. J., de Rooij, R. J., van Boom, J. H., & Altona, C. (1980) *Nucleic Acids Res. 8*, 169–181.

Haasnoot, C. A. G., DeBruin, S. H., Berendsen, R. G., Janssen, H. G. J. M., Binnendijk, T. J. J., Hilbers, C. W., Van der Marel, G. A., & Van Boom, J. H. (1983) *J. Biomol. Struct. Dyn. 1*, 115–129.

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

Hakimelahi, G. H., Proba, Z. A., & Ogilvie, K. K. (1982) *Can. J. Chem. 60*, 1106–1113.

Holbrook, S. R., Sussman, J. L., Warrant, R. W., & Kim, S.-H. (1978) *J. Mol. Biol. 123*, 631–660.

Ikuta, S., Chattopahyana, R., Ito, H., Dickerson, R. E., & Kearns, D. R. (1988) *Biochemistry 25*, 4840–4849.

Johnson, P. D., & Redfield, A. G. (1978) *Nucleic Acids Res. 5*, 3913–3927.

Johnson, P. D., & Redfield, A. G. (1981) *Biochemistry 20*, 1147–1156.

Krol, A., Westhof, E., Bach, M., Luhrmann, R., Ebel, J.-P., & Carbon, P. (1990) *Nucleic Acids Res. 18*, 3803–3811.

McConnell, B., & Seawell, P. C. (1972) *Biochemistry 11*, 4382–4392.

Mount, S. M. (1982) *Nucleic Acids Res. 10*, 459–472.

Ogilvie, K. K., Schifman, A. L., & Penney, C. L. (1979) *Can. J. Chem. 57*, 2230-2235.
Petersheim, M., & Turner, D. H. (1983) *Biochemistry 22*, 256-263.
Pochon, F., Michelson, A. M., Grunberg-Manago, M., Cohn, W. E., & London, L. (1964) *Biochim. Biophys. Acta 80*, 441-447.
Puglisi, J. D., Wyatt, J. R., & Tinoco, I., Jr. (1990) *Biochemistry 29*, 4215-4226.
Reddy, R. (1989) *Methods Enzymol. 180*, 521-532.
Redfield, A. G., Kunz, S. D., & Ralph, E. K. (1975) *J. Magn. Reson. 19*, 114-117.
Richards, E. G. (1975) in *Handbook of Biochemistry and Molecular Biology*: *Nucleic Acids*, 3rd ed. (Fasman, G. D., Ed.) Vol. I, p 197, CRC Press, Cleveland, OH.
Seraphin, B., Kretzner, L., & Roshbash, M. (1988) *EMBO J. 7*, 2533-2538.
Shapiro, R., & Chambers, R. W. (1961) *J. Am. Chem. Soc. 83*, 3920-3921.
Siliciano, P. G., & Guthrie, C. (1988) *Genes Dev. 2*, 1258-1267.
Sprinzl, M., Hartmann, T., Weber, J., Blank, J., & Zeidler, R. (1989) *Nucleic Acids Res. 17* (Suppl), 1-172.
Sung, W. L., & Narang, S. A. (1982) *Can. J. Chem. 60*, 111-120.
Turner, D. H., Sugimoto, N., & Freier, S. M. (1988) *Annu. Rev. Biophys. Biophys. Chem. 17*, 167-192.
Usman, N., Ogilvie, K. K., Jiang, M.-Y., & Cedegren, R. J. (1987) *J. Am. Chem. Soc. 109*, 1845-1854.
Varani, G., Wimberly, B., & Tinoco, I., Jr. (1989) *Biochemistry 28*, 7760-7772.
Wuthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, John Wiley and Sons, New York.
Zhuang, Y., & Weiner, A. M. (1986) *Cell 46*, 827-835.

Requirement for the β,γ-Pyrophosphate Bond of ATP in a Stage between Transcription Initiation and Elongation by *Escherichia coli* RNA Polymerase†

Miki Fujioka,‡ Tatsumi Hirata,§ and Nobuo Shimamoto*

DNA Research Center, National Institute of Genetics, Mishima, Shizuoka-ken 411, Japan

Received August 14, 1990; Revised Manuscript Received November 6, 1990

ABSTRACT: A linear fragment of DNA was fixed to acrylamide or agarose beads by its ends. When a fragment containing the λ P_R promoter is immobilized and transcribed, the RNA products are unchanged from those obtained on the unfixed DNA. Transcription from the immobilized fragment can be interrupted by diluting the reaction mixture into a large volume of the same buffer. Brief centrifugation allows isolation of the transcription complex with the immobilized DNA. If interruption occurs during elongation, the elongation can be resumed upon a second addition of substrates. If ATP is replaced by a β,γ-unhydrolyzable analogue in the second addition, the elongated products are similar to those obtained when the substrate contain ATP. When ATP is replaced by the analogue at the initiation step, however, the yield of elongated products is decreased to less than one-sixth and that of short abortive products is increased. Thus the ATP analogues are good substrates once elongation has been established in the presence of ATP, but not good enough to get past a stage just after initiation in the absence of ATP. We conclude that the β,γ-pyrophosphate bond of ATP is important for preparation of efficient elongation.

Transcription of genes is mediated by a large protein·DNA complex, and the DNA-dependent RNA polymerase is the major functional component. To investigate the relationship between the structure of the transcription complex and the RNA product, one has to freeze the structure and stop elongation simultaneously. Conventional methods using chelating reagents, substrate analogues, or truncated templates are not always satisfactory. The addition of EDTA stops not only RNA synthesis but also every reaction requiring magnesium ion. When 3′-deoxy substrates or truncated DNA templates are used, the rapid and simultaneous cessation of RNA chain elongation is impossible, and it thus complicates the kinetic interpretation of the results.

To overcome this difficulty, we have devised a novel kinetic method, in which transcription is directed from a linear template DNA fixed to resin beads by its ends, and the transcription complex can be isolated by rapid dilution and centrifugation. Elongation of the RNA chain can thus be interrupted without changing ionic conditions and can be resumed with different substrates. A DNA fragment is easily labeled with biotin (Langer et al., 1981), and thus labeled fragment can be strongly bound to avidin linked to a gel matrix. Such an immobilized DNA fragment was used to isolate preinitiation complex in HeLa nuclear extract where the requirements for ATP in preinitiation and initiation were studied (Arias & Dynan, 1989).

A nascent RNA chain is elongated by the addition of a nucleoside monophosphate to its 3′-end, and α,β-pyrophosphate bonds in substrates are absolutely required. On the other hand,

† This work was supported by Research Grants to N.S. from Nissan Foundation and Ministry of Education.
* To whom correspondence should be addressed.
‡ Present address: Department of Biochemistry and Molecular Biology, Harvard University, 7 Divinity Ave., Cambridge, MA 02138.
§ Present address: Department of Biophysics, Faculty of Science, Nagoya University, Chigusa-ku, Nagoya 464, Japan.

0006-2960/91/0430-1801$02.50/0