Structural features of an influenza virus promoter and their implications for viral RNA synthesis

Sung-Hun Bae*, Hae-Kap Cheong*†, Joon-Hwa Lee*, Chaejoon Cheong†, Masatsune Kainosho‡, and Byong-Seok Choi*§

*Department of Chemistry and National Creative Research Initiative Center, Korea Advanced Institute of Science and Technology, 373-1 Kusung-dong, Yusung-gu, Taejon 305-701, Korea; †Magnetic Resonance Team, Korea Basic Science Institute, 52 Eoun-dong, Yusung-gu, Taejon 305-333, Korea; and ‡Department of Chemistry, Faculty of Science, Tokyo Metropolitan University, 1-1 Minami-ohsawa, Hachioji, Tokyo 192-0397, Japan

Edited by Peter Palese, Mount Sinai School of Medicine, New York, NY, and approved July 10, 2001 (received for review May 28, 2001)

The influenza A virus, a severe pandemic pathogen, has a segmented RNA genome consisting of eight single-stranded RNA molecules. The 5′ and 3′ ends of each RNA segment recognized by the influenza A virus RNA-dependent RNA polymerase direct both transcription and replication of the virus's RNA genome. Promoter binding by the viral RNA polymerase and formation of an active open complex are prerequisites for viral replication and proliferation. Here we describe the solution structure of this promoter as solved by multidimensional, heteronuclear magnetic resonance spectroscopy. Our studies show that the viral promoter has a significant dynamic nature and reveal an unusual displacement of an adenosine that forms a novel (A-A)·U motif and a C-A mismatch stacked in a helix. The characterized structural features of the promoter imply that the specificity of polymerase binding results from an internal RNA loop. In addition, an unexpected bending (46 ± 10°) near the initiation site suggests the existence of a promoter recognition mechanism similar to that of DNA-dependent RNA polymerase and a possible regulatory function for the terminal structure during open complex formation.

The influenza A virus, a member of the Orthomyxoviridae family, has a genome consisting of eight single-stranded RNA molecules of negative polarity. These eight RNA segments encode ten proteins, including three RNA-dependent RNA polymerase (RdRp) proteins (PA, PB1, PB2) and one nucleoprotein (NP). The transcription and replication of the viral genome are performed in the nucleus of infected cells by a ribonucleoprotein (RNP) complex that is composed of the three polymerase proteins, NP, and the viral RNA (vRNA). Replication of the vRNA produces a full-length copy of positive-sense RNA (cRNA). The cRNA is then used in the formation of another RNP complex, which serves to generate newly synthesized vRNA. Transcription is also performed by the same RNP complex; however, initiation occurs by the pirating of a 7-methyl guanosine cap structure from a host mRNA. Viral transcription produces an mRNA molecule that is 15 to 22 nucleotides (nt) shorter than cRNA and contains a poly(A) tail (1).

The influenza A virus RdRp binds specifically to the partial duplex promoter (also referred to as the panhandle RNA), which is made from the 5′ and 3′ termini of each RNA genome segment. Within this partial duplex, 13 nt at the 5′ end and 12 nt at the 3′ end are highly conserved among most influenza A virus variants (Fig. 1*A*). All of the necessary signals for replication and genome packaging appear to reside in these terminal sequences (2), and several pieces of evidence imply a regulatory role for the terminal structure in viral transcription initiation (3–6), termination, and polyadenylation (7, 8).

The RdRps are for the most part virus-encoded polymerases that synthesize RNA from an RNA template without a DNA intermediate. Despite the unique features of RdRps, little has been deciphered about their mechanisms of promoter recognition and RNA synthesis. The promoter of the influenza A virus's RNA genome consists of a partial duplex that is similar to a DNA promoter. Therefore, when the promoter–RdRp interaction of the influenza A virus is investigated, comparisons can be made with the large body of information that exists for DNA-dependent RNA polymerase (DdRp)–promoter interactions.

In this study, we determined the solution structure of the 31-mer RNA oligonucleotide containing the conserved influenza A virus promoter sequence. It reveals a structural basis for the requirement of the internal loop for specific interaction with the influenza A virus RdRp. Furthermore, the structure reveals an unexpected bending near the transcription initiation site, which is the first entity of a promoter recognized by a RdRp. The presence of these structural features suggests that RdRps initiate RNA synthesis by a mechanism that is similar to that of DdRps.

Methods

RNA Sample Preparation. The RNA samples were prepared by cleaving a substrate RNA (41 nt) with a trans-cleaving hammerhead ribozyme RNA (44 nt). The substrate RNA and the ribozyme RNA were prepared enzymatically from a synthetic DNA template by using T7 RNA polymerase and either unlabeled rNTPs or ^{13}C,^{15}N-labeled rNTPs (9, 10). Each transcribed RNA was purified by Prepcell (Bio-Rad). The trans-cleavage reaction was performed at 55°C for 60 min with a substrate to ribozyme molar ratio of 10:1 in 50 mM Tris (pH 7.9)/28 mM $MgCl_2$. The cleaved RNA product (31 nt and 10 nt) was gel purified on 15% polyacrylamide/7 M urea gels, because the Prepcell had only about 5-nt resolution. NMR samples were 0.5 mM for a fully ^{13}C,^{15}N-labeled RNA strand and 2 mM for an unlabeled RNA strand in 0.01 mM EDTA/10 mM sodium phosphate (pH 6.5). Purple membrane (PM, Munich Innovative Biomaterials, Marburg, Germany) was added to 2 mg/ml for the measurement of residual dipolar couplings (11, 12).

NMR Spectroscopy. NMR spectra were recorded on Bruker DRX 400, 600 and 800 MHz spectrometers, processed by using XWINNMR 2.6 (Bruker, Billerica, MA), and analyzed by using SPARKY 3.87 (13). Two-dimensional (2D) nuclear Overhauser effect spectroscopy (NOESY) spectrum in 90% H_2O/10% D_2O with a mixing time of 100 and 200 ms were recorded on the unlabeled RNA sample at 278 K. A 2D ^{1}H-^{15}N heteronuclear single quantum correlation (HSQC) spectrum was recorded on the ^{13}C,^{15}N fully labeled RNA sample. The other spectra were all obtained in 99.96% D_2O. Double quantum filtered correlated spectroscopy (DQF-COSY) spectra were recorded at 272 K and 303 K and homonuclear total correlation spectroscopy (TOCSY)

This paper was submitted directly (Track II) to the PNAS office.

Abbreviations: RdRp, RNA-dependent RNA polymerase; DdRp, DNA-dependent RNA polymerase; nt, nucleotide; NOESY, nuclear Overhauser effect spectroscopy; 2D, two-dimensional; DQF-COSY, Double quantum filtered correlated spectroscopy; TOCSY, total correlation spectroscopy; HCCH, proton-carbon-carbon-proton; HSQC, heteronuclear single quantum correlation spectroscopy.

Data deposition: The atomic coordinates have been deposited in the Protein Data Bank, www.rcsb.org (PDB ID code 1JO7).

§To whom reprint requests should be addressed. E-mail: bschoi@cais.kaist.ac.kr.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

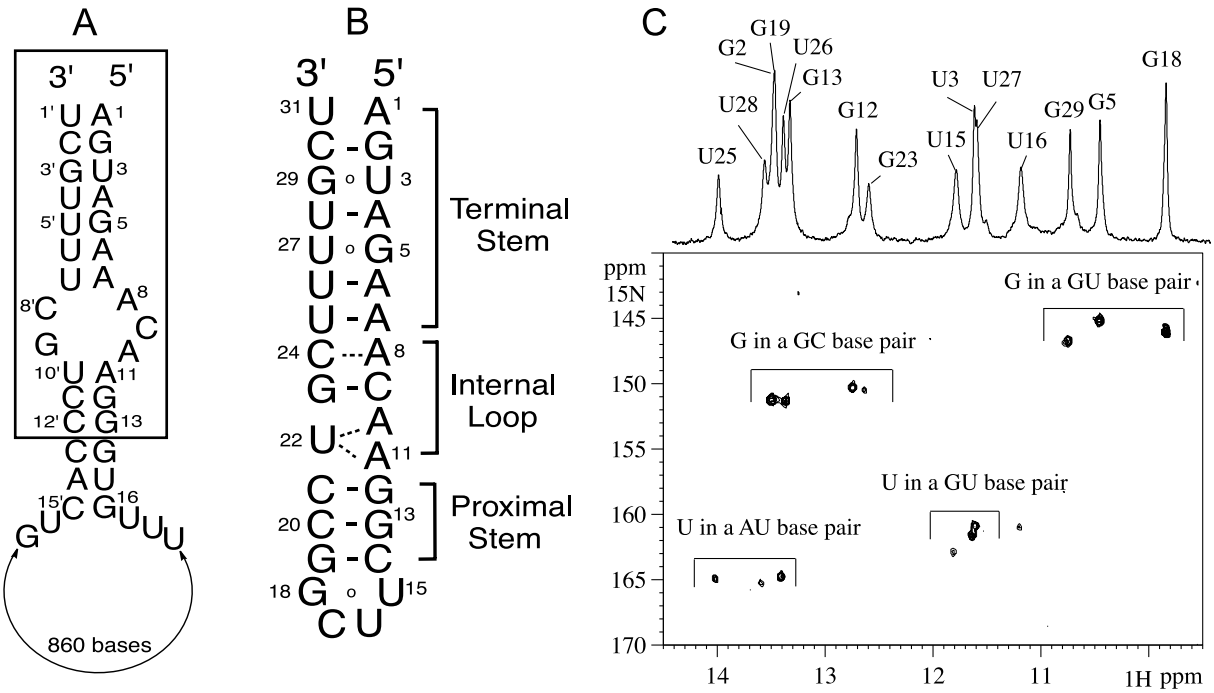

Fig. 1. (*A*) Terminal sequence of the influenza A virus genomic RNA. The conserved 13 nt at the 5′ end and 12 nt at the 3′ ends are boxed. The numbering from the 3′ end is followed by a prime notation (′). The sequence shown is that of viral RNA (vRNA) segment 8 of influenza A/PR/8/34. (*B*) Secondary structure of the RNA construct used in this study. Watson–Crick and non-Watson–Crick base pairs are distinguished by bars and circles, respectively. Dashed lines indicate the calculation-driven base pairs. (*C*) Imino region of ^{1}H-^{15}N HSQC spectrum at 278 K.

spectrum was recorded with a mixing time of 125 ms. Proton-detected ^{1}H-^{31}P heteronuclear correlation (HETCOR; ref. 14) and ^{1}H-^{31}P Hetero-TOCSY with a mixing time of 30 ms were recorded at 297 K. Two-dimensional NOESY spectra with mixing times of 80, 150, and 250 ms were obtained at 294 K, and the residual HDO resonance was suppressed by using low-power presaturation. Two-dimensional ^{1}H-^{13}C constant time HSQC, 2D proton-carbon-carbon-proton (HCCH)-COSY, 2D HCCH–relayed COSY(RELAY), 2D HCCH-TOCSY, and three-dimensional (3D) HCCH-TOCSY (15) with a mixing time of 15 ms were recorded at 294 K.

NMR Assignments and Structural Calculation. All of the slowly exchanging imino and amino protons were assigned by using H_2O NOESY spectra and were confirmed by a ^{1}H-^{15}N HSQC. All of the base protons, H1′, and H2′, most of H3′ and H4′, and several H5′/H5″ resonances were assigned in D_2O experiments including 2D NOESY, DQF-COSY, TOCSY, ^{1}H-^{13}C HSQC, and 3D HCCH-TOCSY. A total of 563 NOE-derived distance constraints were used for the structure calculations. The distance constraints obtained in D_2O NOESY experiments with 80-, 150-, and 250-ms mixing times were grouped into four categories: 1.8–3.4 Å, 2.5–4.5 Å, 3.5–6.0 Å, and 4.0–7.0 Å. The H_2O NOESY-derived distance constraints were grouped into three categories: 1.8–3.4 Å, 1.8–5.0 Å, and 3.8–7.0 Å. No hydrogen bonding constraints were used for the non-Watson–Crick base pairs. The δ dihedral angles were derived from the analysis of $^{3}J_{H1',H2'}$ in DQF-COSY spectra. All χ dihedral angles were constrained to $-158 \pm 15°$ based on the medium to weak intraresidue H1′-H6/H8 NOE, except for the G18 (unconstrained). Backbone dihedral angles (α, β, γ, ε, and ζ) were constrained to the standard A form in the stem regions, which show normal chemical shifts and coupling constants. None of the backbone dihedral angles were constrained in the bulge regions except that a few β and ε angles derived from ^{1}H-^{31}P HETCOR were constrained. Scalar $^{1}J_{CH}$ and dipolar $^{1}D_{CH}$ couplings ($^{1}J_{CH\ anisotropic} - {}^{1}J_{CH\ isotropic}$; see Table 2, which is published as supporting information on the PNAS web site, www.pnas.org) were derived from sensitivity-enhanced HSQC experiments. The precision of the observed dipolar couplings was ± 1 Hz. Alignment tensor analysis of the observed residual dipolar couplings was performed by using single-value decomposition (16, 17). The anisotropy and rhombicity required in the structure refinement were estimated to be -8.0 Hz and 0.32 (see Fig. 5, which is published as supporting information on the PNAS web site), respectively (17, 18). All structure calculations were performed with X-PLOR 3.1 and CNS (19). One hundred starting structures were generated with distance geometry using full structure embedding. The structures were then subjected to a simulated annealing protocol of 10 ps at 3,000 K, followed by 50 ps of cooling to 300 K. The distance force constant was 50 kcal mol^{-1}·$Å^{-2}$, and the dihedral angle constant varied from 20 to 400 kcal mol^{-1}·rad^{-2} when cooling. The lowest energy structures were subjected to a final 20-ps refinement step at 300 K with a time step of 0.5 fs followed by 2,000 cycles of energy minimization. The last 5 ps of each run were averaged and subjected to restrained energy minimization. Two hundred-twenty structures derived from 11 DGSA ensemble structures were used to initiate the refinement protocol with 22 additional residual dipolar coupling restraints that did not belong to the highly dynamic residues. The force constant for residual dipolar coupling energy term was set to 3.0 kcal/mol. The Final 32 structures were then analyzed by Insight II (Biosym Technologies, San Diego) and CURVES 5.2 (20).

Results

Sample Preparation by Means of Trans Cleavage Reaction. During preparation of the 31-mer RNA oligonucleotide (Fig. 1*B*), the *in vitro* transcription reaction using T7 RNA polymerase resulted in negligible transcriptional yield. This was because *in vitro* transcription strictly requires at least one or two G residues at the 5′ end (9) and no U residues within the + 1 to + 6 region of

Table 1. Structure determination statistics for the 32 converged structures of the influenza A virus promoter

Number of NOE distance restraints	563
Internucleotide	278
Intranucleotide	285
Internal loop (residues 8–11, 22–24)	93
Terminal stem (residues 1–7, 25–31)	252
Dihedral restraints (α, β, γ, δ, ε, ζ and χ)	169
Base pair restraints including hydrogen bonding	68
Base planarity restraints	31
Residual dipolar coupling restraints	22
Total number of restraints	853
rmsd for all heavy atoms relative to the mean structure (Å)	
Terminal stem (residues 1–7, 25–31)	0.7 ± 0.4
Internal loop (residues 8–11, 22–24)	1.6 ± 0.1
UUCG tetraloop (residues 14–19)	0.4 ± 0.1
All nucleotides	2.5 ± 0.6
Average NOE violations, Å	0 (>0.5 Å)
Average dihedral angle violations (°)	0 (>5°)
Mean deviation from covalent geometry	
Bond lengths, Å	0.002
Angles, degrees	0.6
Impropers, degrees	0.3

rmsd, root-mean-square deviation.

transcript sequences (H.-K.C., unpublished data). These requirements cannot be met with the conserved 5′ end sequence of the influenza A virus promoter. Therefore, we designed an efficient trans-cleaving hammerhead ribozyme (see Fig. 6, which is published as supporting information on the PNAS web site), which enabled the preparation of an RNA molecule with the wild-type sequence intact. The trans-cleavage reaction was very specific and efficient within a broad range of reaction conditions. The designed trans-cleaving ribozyme had a high turnover rate, such that the cleavage efficiency was similar up to a substrate to ribozyme molar ratio of 50:1 at 55°C. Trans- or cis-cleaving ribozymes have been used mainly for the removal of 3′ heterogeneity of the T7 RNA polymerase transcripts. Recently, however, 5′ heterogeneity (in addition to the well known 3′ heterogeneity) was reported for T7 RNA polymerase transcripts (21). Therefore, the method used in this study should be useful to prepare large quantities of RNA with an arbitrarily defined and homogeneous 5′ end sequence for structural studies.

Overall Conformation. H_2O NOESY and ^1H-^{15}N HSQC spectra confirmed the formation of the expected six and three base pairs in the terminal and proximal stems, respectively. Also, NOE connectivity in D_2O NOESY spectra showed that both stems have the expected A-form geometry. The internal loop showed only one imino resonance of G23, which shows a typical NOE pattern of Watson–Crick base pair with C9. Additional imino resonance did not appear even at pH values as low as 5.5 or on addition of Mg^{2+} ions up to 5 mM (data not shown). Total 853 constraints derived from NMR data were used for structure calculations to obtain the 32 converged structures (Table 1). The average structure is shown in Fig. 2*A*. The structure consists of terminal stem, internal loop, proximal stem, and UUCG tetraloop, which is similar to the previously solved tetraloop structure (22).

Internal Loop. The structure of the internal loop consists of a novel (A-A)·U motif and a C-A mismatch stacked in a helix. The extension of a helix pattern into the internal loop resulted primarily from the stacking of C24 in the helix (Fig. 2*A*). C24 has sequential $H6_i$ to $H1'_{i-1}$, $H6_i$ to $H2'_{i-1}$, and $H6_i$ to $H3'_{i-1}$ nuclear Overhauser effects (NOEs), as well as the sequential $H6_i$ to $H5_{i+1}$ NOE, indicating the continuous stacking of G23, C24, and U25. One of the amino protons of C24 and N1 of A8 exhibited hydrogen bond geometry and the existence of this potential hydrogen bond was partially supported by the observation of different chemical shifts for the two C24 amino protons (see Fig. 2*B*). The $A8^+$·C24 base pair in which the N1 of adenine is protonated and its typical chemical shift is about 14.5 ppm (23) was not observed. The stacking of A8·C24 mismatch contributes to the formation of the C9·G23 base pair and consequently stabilizes the internal loop. The stacked structure of C9 and C24 is quite consistent with the results showing that the N3 of both C9 and C24 are protected from dimethyl sulfate modification (24). The C9 unusually showed no H5–H6 cross peak in the DQF-COSY spectra that were obtained at both 30°C (Fig. 3*A*) and −1°C (data not shown). This feature, together with the significant field-dependent broadness of the H5 and H6 resonances of C9 (Fig. 3*B*), implies that the C9 and/or its neighbor can be involved in dynamic motion on a millisecond time scale while maintaining the Watson–Crick base pair with G23.

A10 and A11, like other adenosines, exhibited the weak sequential $H2_i$ to $H1'_{i+1}$ NOEs typical of A-form RNA helices. We also observed a cross-strand NOE between A10 H2 and G23 H1′, indicating that A10 base stacks into the helix on the top of the C9·G23 base pair. However, the A11 H2 also displayed a cross-strand NOE to G23 H1′ (Fig. 3*C*), suggesting that A11 is displaced toward the minor groove. This is an unusually long-range NOE, because this distance should be about 8 Å in the context of a normal A-form helix. Nevertheless, the pattern of intra- and interresidue NOEs (H6/H8-H1′, H6/H8-H2′, and H6/H8-H3′) continued through the 5′ strand of the internal loop (A8-A11). It implies that the stacking interactions are not broken despite the A11 displacement; thus, a kind of helical axis rearrangement should be accompanied by this adenosine displacement. In addition, A10 was involved in interstrand stacking with G23, which resulted in the positioning of A10 H2 over the purine ring of G23. This geometry caused the upfield chemical shift of the A10 H2 (7.05 ppm), due to the ring current effect. The $^3J_{H1'\text{-}H2'}$ coupling constant is a signature of a sugar pucker, and a $^3J_{H1'\text{-}H2'}$ of 5 to ≈6 Hz, which was observed in A10 (5.1 Hz), A11 (6.2Hz), and G23 (6.7 Hz), infers that dynamic equilibrium exists between the C3′ endo ($^3J_{H1'\text{-}H2'} < 2$ Hz) and C2′ endo ($^3J_{H1'\text{-}H2'} > 8$ to ≈9 Hz) sugar conformations in those internal loop residues.

The displacement of A11 into the minor groove results in two remarkable helical features of the internal loop: the major groove becomes wider and deeper, and the helical axis rearranges at the junction of internal loop and proximal stem. These structural features are consistent with the chemical modification result (24) in that the N7 positions of both A10 and A11 are readily accessible by diethylpyrocarbonate.

(A-A)·U motif. The consecutive adenosines (A10–A11) and a uracil (U22) on the opposite strand can form a (5′-A-A-3′)·U motif. This motif is a binding site for ribosomal protein S8 in *Escherichia coli* 16S rRNA (25) and constitutes the site of spliceosome binding to the branch-point helix and the RNA binding site for phage GA coat protein (26). We have not observed the U22 imino resonance even in low pH conditions (pH 5.5), probably because of the rapid solvent exchange. The highly dynamic nature of the internal loop suggests that U22 might exist in several milieus: for example, in equilibrium between the two possible A-U base pairs or in a base-triple interaction. A base-triple interaction of an (A-A)·U motif by protonation of N1 of the 3′ adenosine at pH 6.1 was observed by Smith and Nikonowicz (26). We, however, did not observe such a protonation. Instead our results show that the (A-A)·U motif can be classified into three groups (see Fig. 2*C*). In the preferred conformations (15 of 32), the U22 base was shifted to the major

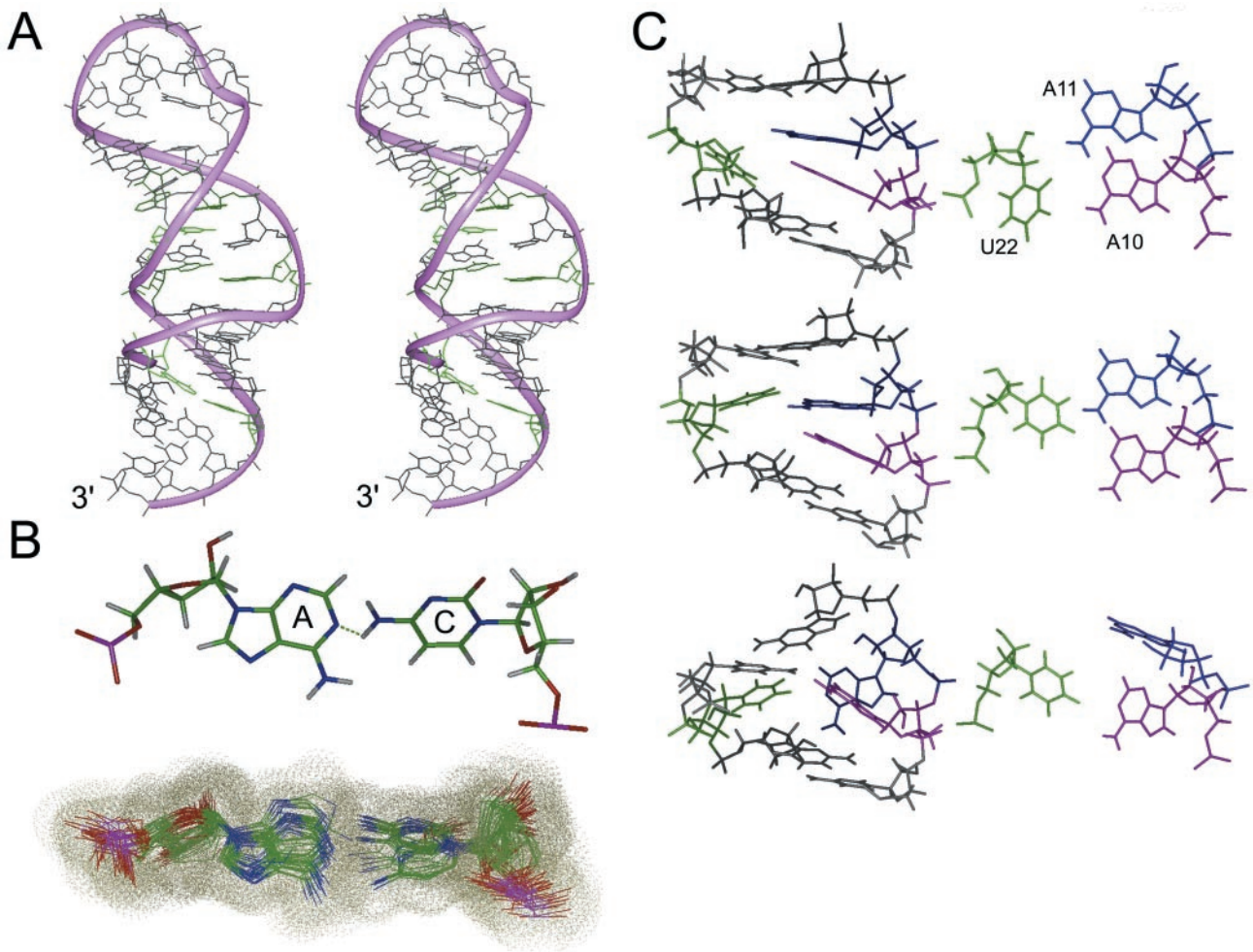

Fig. 2. (*A*) The overall structure of the influenza A virus promoter. A4·U28 in the terminal stem, A8·C24 mismatch, and the (A-A)·U motif in the internal loop are shown in green. (*B*) A potential hydrogen bond formed between A8 and C24 (up) and superposition of 32 structures of the A·C mismatch (down). (*C*) Three groups of low-energy structures of the (A-A)·U motif are shown, viewed from the major groove (*Left*) and from the top (*Right*). A10 is shown in purple, A11 in blue, and U22 in green.

groove whereas the A11 base was shifted to minor groove (Fig. 2*C Top*). In the second most likely conformations (9 of 32), the U22 and A11 bases assumed a rather coplanar geometry and A10 was stacked between A11 and the C9-G23 base pair (Fig. 2*C Middle*). In the third most likely conformations (8 of 32), A11 bulged out of the helix and its normal of the base plane was nearly perpendicular to the helical axis (Fig. 2*C Bottom*). None of these conformations displayed the specified hydrogen bonding geometry. Given the available structural information, we concluded that none of these conformations could be ignored and, rather, that the nature of the internal loop can be explained by the merging of these conformations.

The extensive mutations in A11 and/or U22 severely reduced the promoter efficiency to 2–20% of that of the wild-type promoter. Complementary mutations that restored the base pair between the 10 or 11 and 22 positions did not reestablish promoter efficiency (27). This result and our structural data suggest that A10 or A11 and U22 do not form a Watson–Crick base pair. Instead, these adenosines present specific functional groups (such as the amino group) in the major groove for interaction with the viral polymerase.

Terminal Stem. The terminal stem is well converged with a root-mean-square deviation (rmsd) of 0.7 Å with respect to the average structure (Fig. 4*A*). The helical parameter analysis revealed that the global helical axis of the terminal stem is bent toward the major groove at the A4·U28 step with an angle of 46 ± 10° (Fig. 4*B*). Despite the presence of the bending, U28 stacked on U27 and A4 stacked on G29, resulting in a large propeller twist angle (−23°) for the A4·U28 base pair; this propeller twist angle was much larger than that of a normal A-form helix (−14°). This asymmetric helical feature was also evident from the absence of the imino to imino NOEs of U28-U3 and U28-G29, which was contrary to the presence of the imino to imino NOEs of U28-G5 and U28-U27 in the H_2O NOESY spectra. It is possible that the bending is accommodated by the local dynamics of the U28 and A6 residues, as evidenced by their $^3J_{H1'\text{-}H2'}$ of 4.8 Hz and 4.9 Hz, respectively.

The bending is accompanied by unusual twist angles of nearby base pairs. From the U3·G29 base pair to the A6·U26 base pair, the twist angles with respect to the previous base pair were 37°, 32°, 31°, and 42°, each of which deviates significantly from the average value of the terminal stem (34°). This under- and over-twist angle pattern of the neighboring base pairs of the G·U wobble pair also has been found in several crystal structures containing G·U wobble pairs (28).

In a previously solved structure of the influenza A virus promoter (29), the bending was not observed. This difference is

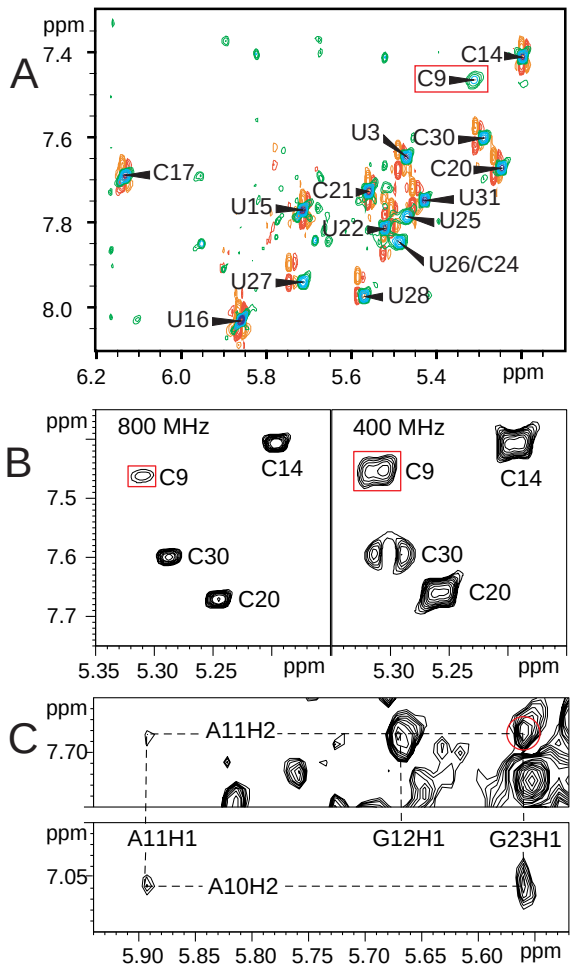

Fig. 3. (*A*) The pyrimidine H5 to H6 region of an 80-ms NOESY spectrum (green-blue) superimposed on the DQF-COSY spectrum (orange-red). Each H5-H6 cross peak of pyrimidines was assigned on the NOESY spectrum. (*B*) The selected NOE cross peaks of H5 and H6 of pyrimidine residues obtained from 400 MHz and 800 MHz NOESY spectra at 294 K. (*C*) The H2 to H1′ region of a 150-ms NOESY spectrum recorded at 294 K. The H2 resonances of A10 and A11 are indicated by a horizontal dashed lines, and the H1′ resonances of A11, G12, and G23 are indicated by a vertical dashed line. The red circled cross peak indicates the unusual NOE between A10H2 and G23H1′.

primarily due to the substitution of the U3 with a cytosine. Two other differences between this and the previous structure (C24 bulge and A11·U22 base pair) might also be a result of this substitution and its consequent effects. These structural differences are based on independently obtained spectral data and could be rationalized by considering long-range effects (30), such as the helical phase differences.

Discussion

Transcription by a DdRp can be separated into a series of steps, including DNA template binding by the DdRp, promoter localization, melting of the DNA to form an active open complex, nucleotide substrate binding, formation of the first phosphodiester bond, abortive RNA synthesis, promoter clearance, transcriptional elongation, and termination (31). Similarly, during the initiation of RNA synthesis by an RdRp using a double-stranded RNA promoter, the polymerase should locate the promoter and arrange its catalytic core at the transcription initiation site and, with the nucleic acid template, form a so-called open complex.

The specific residues of promoter involved in influenza A virus RdRp binding were proposed as A8-A11 and/or C20-G23, as well as U28-U31 (32, 33). Most of these residues belong to the internal loop, which has a unique and highly dynamic structure as described. The importance of the internal loop structure in polymerase binding can be understood in the context of the entropy contribution. If the internal loop is single-stranded and unstructured, randomly positioned and oriented functional groups might be rearranged to a specific configuration by the viral RdRp at the cost of entropy decrement to identify the promoter. Therefore, the polymerase should prefer the preformed internal loop, and the dynamic nature of the internal loop should facilitate the induced fitting mode of protein–RNA interaction.

Also, the bending at the A4·U28 would assist the polymerase to locate the start site of RNA synthesis. As observed with the TATA binding protein (34), the influenza A virus polymerase might anchor at the minor groove of the bending site, because the bending is toward the major groove and leaves the minor groove exposed. The localized dynamic properties of U28 and A6 also imply that the potential polymerase contacts occur at the bending site. Thus, we propose that terminal stem structure has a critical role during promoter localization, even though the major determinant of the specific recognition of the influenza A virus promoter has been known to reside in the internal loop.

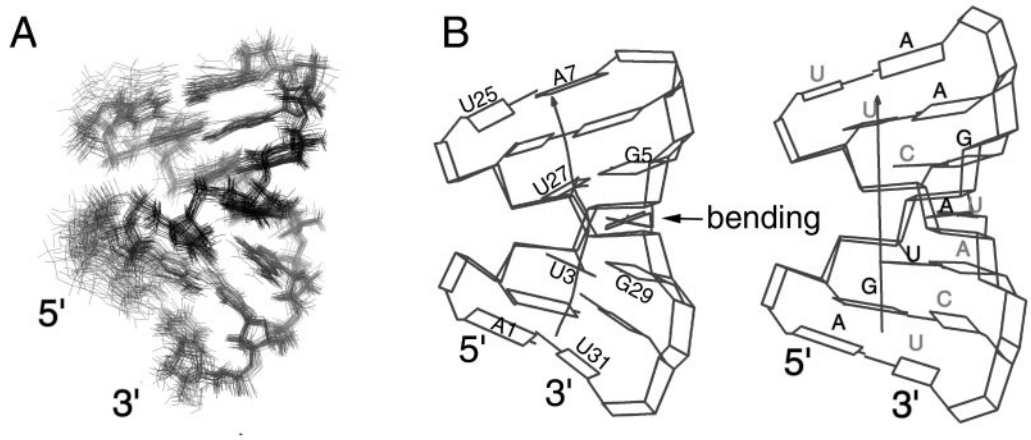

Fig. 4. (*A*) The 32 converged structures, superimposed on the terminal stem. (*B*) Schematic representation of the helical geometry of the terminal stem (*Left*) and of a normal A-form helix (*Right*) with a complete Watson–Crick base pair (5′-AGUAGAA/3′-UCAUCUU). The one-headed arrow indicates the helical axis. This figure was generated with CURVES 5.2.

The boundary of the promoter residues involved in open complex formation of influenza A virus transcription was inferred from the analysis of abortive transcripts and the NTP requirement during the initiation of RNA synthesis (35). The results suggest that the open complex should extend to the +4 position (A4·U28) and the incorporation of AMP at the +5 position induce the RdRp to carry out transcriptional elongation. Interestingly, this hypothesized boundary of the open complex coincides with the bending site. The bending at the +4 position would lower the energy for unwinding and opening of the helix, which should be achieved before inserting the first nucleotide. For DNA, bending has been known to lower the energy barrier to the melting of the DNA strands (36) and facilitate base pair opening (37). The bacteriophage T7 DdRp is known to bind to and bend its cognate promoter DNA, which is inherently bent (about 10°; ref. 38). The intrinsic bending in the influenza A virus promoter might be even more important, because no helicase activity has been detected for the influenza A virus RdRp, although some viral RdRps do display helicase activity (39).

The 5′ and 3′ ends of the influenza A virus promoter are nearly complementary. However, the existence or the importance of the double stranded terminal structure has been questioned (3, 4). Recently it has been shown that both the 5′ and 3′ end regions of the influenza A virus promoter can bind the influenza A virus RdRp independently, but the binding affinity increases significantly when both the 5′ and 3′ sequences exist as a partial duplex (40). Also, when the influenza A virus promoter was cloned into a more internally located position of a double-stranded RNA, while leaving the terminal-stem-disrupting mutant promoter in the original position, it still retained its polymerase binding function. In addition, the initiation sites of positive-sense RNA (cRNA) and mRNA synthesis were identical to those expected for the promoter in an internally displaced location (41). From these results and our structural data, we infer that the influenza A virus promoter is a double-stranded helical structure and its helical structure has a possible regulatory role in the initiation of RNA synthesis. This appears to be the case even though the viral promoter can be easily melted because of the inherent bending of the terminal stem and the solvent accessibility of the dynamic internal loop.

In summary, our structural results and careful consideration of previously acquired biochemical data provide new insight into the initial steps of RNA synthesis by the influenza A virus RdRp. Specifically, the structural and dynamic features of the internal loop and the bending of the terminal stem assist the viral RdRp in its identification of the start site of RNA synthesis. In addition, the bending property of the terminal stem may facilitate melting of the RNA strands to form an open complex.

We thank Miss Mi-Kyung Lee, Miss Chin-Ju Park, Mr. Jun-Seok Yoo, Mrs. Eun-Hee Kim, and Dr. Chojiro Kojima for the sample preparation and assistance with the NMR experiments. This work was supported by the Molecular Medicine Research Group Program (grant to B.-S.C.), and by an International Research Collaboration Grant (99-I-01-03-A-043 to C.C.) from the Ministry of Science and Technology, the Republic of Korea. S.-H.B. was supported partially by the BK21 project.

1. Lamb, R. A. & Choppin, P. W. (1983) *Annu. Rev. Biochem.* **52,** 467–506.
2. Luytjes, W., Krystal, M., Enami, M., Pavin J. D. & Palese, P. (1989) *Cell* **59,** 1107–1113.
3. Fodor, E., Pritlove, D. C. & Brownlee, G. G. (1994) *J. Virol.* **68,** 4092–4096.
4. Fodor, E., Pritlove, D. C. & Brownlee, G. G. (1995) *J. Virol.* **69,** 4012–4019.
5. Cianci, C., Tiley, L. & Krystal, M. (1995) *J. Virol.* **69,** 3995–3999.
6. Hagen, M., Chung, T. D. Y., Butcher, J. A. & Krystal, M. (1994) *J. Virol.* **68,** 1509–1515.
7. Poon, L. L. M., Pritlove, D. C., Sharps, J. & Brownlee, G. G. (1998) *J. Virol.* **72,** 8214–8219.
8. Pritlove, D. C., Poon, L. L. M., Fodor, E., Sharps, J. & Brownlee, G. G. (1998) *J. Virol.* **72,** 1280–1286.
9. Milligan, J. F., Groebe, D. R., Witherell, G. W. & Uhlenbeck, O. C. (1987) *Nucleic Acids Res.* **15,** 8783–8798.
10. Batey, R. T., Inada, M., Kujawinski, E., Puglisi, J. D. & Williamson, J. R. (1992) *Nucleic Acids Res.* **20,** 4515–4523.
11. Tjandra, N. & Bax, A. (1997) *Science* **278,** 1111–1114.
12. Sass, J., Cordier, F., Hoffmann, A., Rogowski, M., Cousin, A., Omichinski, J. G., Lowen, H. & Grzesiek, S. (1999) *J. Am. Chem. Soc.* **121,** 2047–2055.
13. Goddard, T. D. & Kneller, D. G. (2000) SPARKY3 (University of California, San Francisco).
14. Sklenář, V., Miyashiro, H., Zon, G., Miles, H. T. & Bax, A. (1986) *FEBS Lett.* **208,** 94–98.
15. Kay, L. E., Xu, G.-Y., Singer, A. U., Muhandiram, D. R. & Forman-kay, J. D. (1993) *J. Magn. Reson. B* **101,** 333–337.
16. Losonczi, J. A., Andrec, M., Fisher, M. W. F. & Prestegard, J. H. (1999) *J. Magn. Reson.* **138,** 334–342.
17. Zweckstetter, M. & Bax, A. (2000) *J. Am. Chem. Soc.* **122,** 3791–3792.
18. Clore, G. M., Gronenborn, A. M. & Tjandra, N. (1998) *J. Magn. Reson.* **131,** 159–162.
19. Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, M., Pannu, N. S., *et al*. (1998) *Acta Crystallogr. D* **54,** 905–921.
20. Lavery, R. & Skelenar, H. (1997) CURVES 5.2.
21. Helm, M., Brulé, H., Giegé, R. & Florentz, C. (1999) *RNA* **5,** 618–621.
22. Allain, F.H.-T. & Varani, G. (1995) *J. Mol. Biol.* **250,** 333–353.
23. Puglisi, J. D., Wyatt, J. R. & Tinoco, I., Jr. (1990) *Biochemistry* **29,** 4215–4226.
24. Baudin, F., Bach, C., Cusack, S. & Ruigrok, R. W. H. (1994) *EMBO J.* **13,** 3158–3165.
25. Kalurachchi, K. & Nikonowicz, E. P. (1998) *J. Mol. Biol.* **280,** 639–654.
26. Smith, J. S. & Nikonowicz, E. P. (1998) *Biochemistry* **37,** 13486–13498.
27. Kim, H.-J., Fodor, E., Brownlee, G. G. & Seong, B. L. (1997) *J. Ger. Virol*. **78,** 353–357.
28. Masquida, B. & Westhof, E. (2000) *RNA* **6,** 9–15.
29. Cheong, H.-K., Cheong, C., Lee, Y.-S., Seong, B. L. & Choi, B.-S. (1999) *Nucleic Acids Res.* **27,** 1392–1397.
30. Wimberly, B., Varani, G. & Tinoco, I., Jr. (1993) *Biochemistry* **32,** 1078–1087.
31. McClure, W. R. (1985) *Annu. Rev. Biochem.* **54,** 171–204.
32. Fodor, E., Seong, B. L. & Brownlee, G. G. (1993) *J. Gen. Virol.* **74,** 1327–1333.
33. Tiley, L. S., Hagen, M., Matthews, J. T. & Krystal, M. (1994) *J. Virol.* **68,** 5108–5116.
34. Kim, J. L. & Burley, S. K. (1994) *Nat. Struct. Biol.* **1,** 638–653.
35. Klumpp, K., Ford, M. J. & Ruigrok, R. W. H. (1998) *J. Gen. Virol.* **79,** 1033–1045.
36. Perez-Martin, J. & Espinosa, M. (1994) *J. Mol. Biol.* **241,** 7–17.
37. Ramstein, J. & Lavery, R. (1988) *Proc. Natl. Acad. Sci. USA* **85,** 7231–7235.
38. Ujvari, A. & Martin, C. T. (2000) *J. Mol. Biol.* **295,** 1173–1184.
39. Cho, M. W., Richards, O. C., Dmitrieva, T. M., Agol, V. & Ehrenfeld, E. (1993) *J. Virol.* **67,** 3010–3018.
40. Gonzalez, S. & Ortin, J. (1999) *J. Virol.* **73,** 631–637.
41. Flick, R. & Hobom, G. (1999) *Virology* **262,** 93–103.